

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2002, 17:24:40 ; Search time 12.371 Seconds
(without alignments)
1059.473 Million cell updates/sec

Title: US-09-135-238B-2_COPY_273_390

Perfect score: 616

Sequence: 1 KRAVERKALSRRLAVR.....HQPAAMMEDSDSDYINVPA 118

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_032802.*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	616	100.0	390	20	AA42225 Human Toso protein
2	616	100.0	390	20	AA17496 Human Toso protein
3	616	100.0	390	20	AA105001 Human PIGRL-1 prot
4	356.5	57.9	422	22	AAE05349 Mouse Toso protein
5	93	15.1	407	21	AAAG40735 Zea mays protein f
6	91.5	14.9	538	22	AAAB82806 Human low density
7	90	14.6	416	22	AAU17493 Novel signal trans
8	88	14.3	609	21	AAU53876 A human seven tran
9	87	14.1	314	22	AAU32408 Novel human secret
10	87	14.1	609	22	AAU08993 Human G protein-co
11	87	14.1	609	22	AAAG4292 Human Grp-binding

12	87	14.1	740	22	AAAG3401 Human polypeptide,
13	86	14.0	343	22	ABG12999 Novel human diagno
14	85.5	13.9	2205	16	AAAG79048 Infectious rubella
15	85	13.8	213	22	AAAG90922 C glutamicum prote
16	85	13.8	554	22	ABB71946 Drosophila melanog
17	84.5	13.7	217	22	ABB10356 Human cell surface
18	84.5	13.7	358	21	AAAY94336 Human cDNA SEQ ID
19	84.5	13.7	359	20	AAAY41690 Human PRO329 prote
20	84.5	13.7	359	21	AAAB44246 Human PRO329 (UNQ2
21	84.5	13.7	359	21	AAAB34744 Human secreted pro
22	84.5	13.7	359	22	AAU29044 Human PRO polypept
23	84.5	13.7	359	22	AAE03451 Human gene 25 enco
24	84.5	13.7	550	22	AAAB82807 Rabbit low density
25	84	13.6	466	13	AAAR26505 Adrenalin receptor
26	84	13.6	609	22	AAE04548 Human G-protein co
27	83	13.5	706	22	ABB66693 Drosophila melanog
28	83	13.5	1260	22	ABB64653 Drosophila melanog
29	83	13.5	1740	21	AAAY83017 Rat shank 3a. Rat
30	83	13.5	1740	22	AAAB31517 Amino acid sequenc
31	82.5	13.4	481	22	ABG05355 Novel human diagno
32	82	13.3	381	22	AAAG89886 C glutamicum prote
33	82	13.3	452	22	AAAB94445 Human protein sequ
34	82	13.3	681	22	ABG07912 Novel human diagno
35	82	13.3	802	22	ABB71892 Drosophila melanog
36	82	13.3	981	22	ABG22248 Novel human diagno
37	82	13.3	1115	22	ABG28242 Novel human diagno
38	81.5	13.2	2115	19	AAAG59276 Rubella virus RA27
39	81	13.1	145	22	AAO06808 Human polypeptide
40	81	13.1	194	22	AAAG74948 Human colon cancer
41	81	13.1	292	22	ABG15356 Novel human diagno
42	81	13.1	385	21	AAAB56596 Human prostate can
43	81	13.1	550	16	AAAR72830 Human TIABP2. Hom
44	81	13.1	550	19	AAAW64717 Human TIABP2 prote
45	81	13.1	550	20	AAAY28589 TTA-1 binding prot

ALIGNMENTS

RESULT 1
AA42225
ID AA42225 standard; Protein; 390 AA.
XX
AC AA42225;
XX
DT 17-DEC-1999 (first entry)
XX
DE Human Toso protein sequence.
XX
KW Human; Toso protein; target; drug screening; diagnosis; apoptosis;
XX
OS Homo sapiens.
XX
PN WO9950671-A2.
XX
PD 07-OCT-1999.
XX
PF 30-MAR-1999; 99WO-US06945.
XX
PR 30-MAR-1998; 98US-0050861.
XX
PA (RIGE-) RIGEL PHARM INC.
XX
PI Payan D;
XX
DR WPI: 1999-591379/50.
XX
N-PSDB; AAZ25422.
XX
PT Screening agents useful for modulating apoptosis and controlling
XX
apoptosis related diseases -
XX
Claim 1; Fig 2a; 75pp; English.

XX The present invention describes a method of Screening for a bioactive
 CC agent capable of binding a Toso protein. Also described a methods for:
 CC (1) screening a bioactive agent capable of modulating activity of a
 CC Toso cell-surface receptor, comprising adding a candidate bioactive
 CC agent to a cell comprising a recombinant Toso nucleic acid, and
 CC exposing the cells to an apoptotic agent that will induce apoptosis;
 CC (2) modulating apoptosis comprising administering an exogenous
 CC compound that binds Toso, to a cell; (3) identifying a cell containing
 CC a mutant Toso gene, comprising determining it's sequence; (4)
 CC identifying the Toso genotype, comprising determining the sequence of
 CC at least one Toso gene; and (5) diagnosing an apoptosis related
 CC condition, comprising measuring activity of Toso in a tissue, and
 CC comparing to the activity from non-affected individual's tissue, where
 CC a reduced activity of the patient indicates risk of an apoptosis related
 CC condition. The methods are useful for identifying agents capable of
 CC diagnosing and treating apoptosis related disease, their use for
 CC modulating apoptosis, and methods for diagnosing the disease state.
 CC The present sequence represents the human Toso protein for use in
 CC methods from the invention.
 XX
 SQ Sequence 390 AA;

Query Match 100.0%; Score 616; DB 20; Length 390;
 Best Local Similarity 100.0%; Pred. No. 3.le-53;
 Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRAVERKALSRARRLAVRMRALESSQRPGRSPRSONNIYSACPRRARGADAAGTGE 60
 DB 273 KRAVERKALSRARRLAVRMRALESSQRPGRSPRSONNIYSACPRRARGADAAGTGE 332
 QY 61 APVPGCAPLPAPLOVSESPWLHAPSLKTSCEYVSLYHQPAMMEDSDDDYINVPA 118
 DB 333 APVPGCAPLPAPLOVSESPWLHAPSLKTSCEYVSLYHQPAMMEDSDDDYINVPA 390

RESULT 2
 AAY17496
 ID AAY17496 standard; Protein; 390 AA.
 XX
 AC AAY17496;
 XX
 DT 03-AUG-1999 (first entry)
 XX
 DE Human Toso protein.
 XX
 KW Toso protein; tumour necrosis factor mediated apoptosis inhibition;
 KW TNF mediated apoptosis; T cell overactivity; autoimmune disease;
 KW Sjogrens connective tissue disorder; transplant rejection; cancer.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..17
 FT /label= signal
 FT Protein 18..390
 FT /label= Toso
 FT Region 254..272
 FT /label= transmembrane_region
 XX
 XX WO9925832-A1.
 PN 27-MAY-1999.
 XX
 XX 16-NOV-1998; 98WO-US24391.
 XX
 XX 17-AUG-1998; 98US-0135238.
 PR 17-NOV-1997; 97US-0066063.
 XX
 XX (STRD) UNIV LELAND STANFORD JUNIOR.
 PA
 XX Hitoshi Y, Nolan GP;
 PI
 XX

DR WPI; 1999-338007/28.
 DR N-PSDB; AAX76123.
 XX
 PT DNA encoding Toso, a protein having inhibitory effects on TNF
 PT mediated apoptosis
 XX
 PS Claim 20; Fig 2a; 70pp; English.
 XX
 CC The present sequence is a Toso protein (I). (I) has anti-apoptotic
 CC and cytostatic activity. Toso (named after a Japanese liquor that is
 CC drunk on New Year's Day to celebrate long life and eternal youth) most
 CC likely acts by induction of cFLIP expression which inhibits caspase-8
 CC processing. Recombinant (I) can be used to modulate apoptosis in a cell
 CC or to treat an apoptosis related condition in a mammal. Apoptosis
 CC related conditions can also be treated by administration of the Toso
 CC protein or antibody. Apoptosis related or mediated conditions that can
 CC be treated include diseases characterized by T cell overactivity, e.g.
 CC Sjogrens connective tissue disorder, autoimmune diseases, diseases where
 CC T cells actively destroy cells, including transplant rejection and
 CC conditions where cells of any kind that are not dying express Toso
 CC appropriately, e.g. cancer of T or B cell origin (where increased
 CC apoptosis would be appropriate).
 XX
 SQ Sequence 390 AA;

Query Match 100.0%; Score 616; DB 20; Length 390;
 Best Local Similarity 100.0%; Pred. No. 3.le-53;
 Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRAVERKALSRARRLAVRMRALESSQRPGRSPRSONNIYSACPRRARGADAAGTGE 60
 DB 273 KRAVERKALSRARRLAVRMRALESSQRPGRSPRSONNIYSACPRRARGADAAGTGE 332
 QY 61 APVPGCAPLPAPLOVSESPWLHAPSLKTSCEYVSLYHQPAMMEDSDDDYINVPA 118
 DB 333 APVPGCAPLPAPLOVSESPWLHAPSLKTSCEYVSLYHQPAMMEDSDDDYINVPA 390

RESULT 3
 AAY05001
 ID AAY05001 standard; Protein; 390 AA.
 XX
 AC AAY05001;
 XX
 DT 16-JUN-1999 (first entry)
 XX
 DE Human PIGRL-1 protein sequence.
 XX
 KW PIGRL-1; human; autoimmune disease; hyper-IgM Immunodeficiency; HIM;
 KW X-linked Severe Combined Immunodeficiency; XSCID; Iga deficiency;
 KW diagnosis; therapy.
 XX
 OS Homo sapiens.
 XX
 PN EP905238-A2.
 XX
 PD 31-MAR-1999.
 XX
 PF 14-AUG-1998; 98EP-0306487.
 XX
 PR 30-OCT-1997; 97US-0961564.
 PR 25-AUG-1997; 97US-0056935.
 XX
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 XX
 XX Sweet RW, Truneh A, Wu S;
 XX WPI; 1999-192666/17.
 DR N-PSDB; AAX28178.
 XX
 XX New polypeptides encoding PIGRL-1 useful for treating diseases such
 PT as X-linked Severe Combined Immunodeficiency
 PT
 XX

PS Claim 11; Page 7; 26pp; English.

XX This sequence is the human PIGRL-1 protein of the invention.

CC Autoimmune diseases involving altered expression or activity of PIGRL-1

CC may include Hyper-IgM Immunodeficiency (HIM), X-linked Severe Combined

CC Immunodeficiency (XSCID) and IGA deficiency. These diseases can be

CC diagnosed or susceptibility to them predicted by: (1) determining whether

CC there is a mutation in the genomic copy of the gene encoding PIGRL-1; or

CC (2) measuring the amount of PIGRL-1 in a sample derived from the patient.

CC Patients deficient in PIGRL-1 can be treated by administering either the

CC PIGRL-1 DNA or its complement or an agonist of PIGRL-1 to the patient.

CC Patients with excessive expression or activity of PIGRL-1 can be treated

CC by administering an antagonist of PIGRL-1, an antisense nucleic acid

CC molecule which inhibits the expression of PIGRL-1 or administering

CC sufficient PIGRL-1 to compete with the endogenous activity. PIGRL-1 can

CC be used to identify its agonists by contacting a cell expressing PIGRL-1

CC with a candidate compound in the presence of a signal system and noting

CC the candidate as an agonist if a signal is produced. The same method can

CC be used to identify antagonists of PIGRL-1 but the presence of an

CC antagonist is indicated by a decrease in production of the signal.

CC Antibodies against PIGRL-1 may be used to isolate or identify clones

CC expressing PIGRL-1. Polynucleotides encoding PIGRL-1 may be used to

CC identify chromosomal mutations in the gene encoding PIGRL-1 in patients.

CC This information may then be correlated with the incidence of autoimmune

CC disease in those patients to identify whether the mutation causes the

CC disease.

XX

SQ Sequence 390 AA;

Query Match 100.0%; Score 616; DB 20; Length 390;

Best Local Similarity 100.0%; Pred. No. 3.1e-53;

Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRAVERKALSRARRLAVRMRALESQRPRGSRPRSONNIYSACPRRARGADAAGTGE 60

Db 273 KRAVERKALSRARRLAVRMRALESQRPRGSRPRSONNIYSACPRRARGADAAGTGE 332

Qy 61 APVPGCAPLPAPLOVSESPWLHAPSLKTSCEYVSLYHOPAAHMEDESDDDYNVPA 118

Db 333 APVPGCAPLPAPLOVSESPWLHAPSLKTSCEYVSLYHOPAAHMEDESDDDYNVPA 390

RESULT 4

AAE05349

ID AAE05349 standard; Protein; 422 AA.

XX

AC AAE05349;

XX

DT 12-SEP-2001 (first entry)

XX

DE Mouse Toso protein.

XX

KW Mouse; cytostatic; antiinflammatory; immunoregulatory; tissue integrity;

KW wound healing; immune response; vaccine; cancer; asthma; allergy;

KW cell trafficking; therapy; secreted protein; Fas-induced apoptosis;

Toso.

XX

OS Mus sp.

XX

PN WO200148192-A1.

XX

PD 05-JUL-2001.

XX

PF 21-DEC-2000; 2000WO-NZ00256.

XX

XX 23-DEC-1999; 99US-0171678.

PR 28-NOV-2000; 2000US-0724864.

XX

PA (GENE-) GENESIS RES & DEV CORP LTD.

XX

PI Watson JD, Murison JG;

XX

DR WPI; 2001-425665/45.

DR

DR N-PSDB: AAD10117.

XX Novel isolated polypeptide useful to isolate corresponding interacting

PT proteins or other compounds, to quantitatively determine levels of

PT interacting proteins or other compounds, and as therapeutic target -

XX

PS Claim 6; Page 78-79; 101pp; English.

XX

CC The patent discloses novel polynucleotides and their corresponding

CC proteins which play a major role in induction of growth, cell migration

CC and proliferation, cell-cell interaction and the differentiation of

CC tissue-specific cells. These proteins are important in the maintenance

CC of tissue integrity and thus are important in wound healing. They are

CC useful in various assays to determine the biological activity, to raise

CC antibodies, to isolate corresponding interacting proteins or other

CC compounds, to quantitatively determine levels of interacting proteins or

CC other compounds, and as therapeutic target in a whole range of disease

CC states. Compositions comprising the novel proteins of the invention are

CC useful for treating mammalian disorders. Polynucleotides of the invention

CC are useful in genome and physical mapping, in positional cloning of

CC genes, to tag or identify an organism or its reproductive material (as

CC non-disruptive tags for marking organisms), and for the diagnosis and

CC treatment of mammalian diseases which is the consequence of inappropriate

CC expression of kinase genes. They are useful for promoting immune response

CC as part of a vaccine or anti-cancer treatment, as target for cancer

CC treatment, as immunoregulatory and anti-inflammatory molecule, as

CC diagnostic for specific types of cancer and for development of an

CC anti-cancer treatment, and as a target for antagonists in the treatment

CC of diseases such as asthma and allergy. They are also useful to inhibit

CC or enhance the activity of the soluble molecule that binds proteins of

CC the invention, for tissue and neural regeneration, to promote or block

CC cell trafficking, and as anti-inflammatory and/or vaccine adjuvant.

CC The present sequence is Toso, a secreted protein from mouse. Toso is

CC a cell surface, specific regulator of Fas-induced apoptosis in T-cells.

XX

SQ Sequence 422 AA;

Query Match 57.9%; Score 356.5; DB 22; Length 422;

Best Local Similarity 59.8%; Pred. No. 2.2e-27;

Matches 73; Conservative 12; Mismatches 32; Indels 5; Gaps 1;

Qy 1 KRAVERKALSRARRLAVRMRALES-----QRPRGSRPRSONNIYSACPRRARGADA 55

Db 284 KRAIQRRRRASSRRAGRLAMRRRGRCASRPFTQRRDAPQRPRSONNIYSACPRRARGPDS 343

Qy 56 AGTGEAPVPGCAPLPAPLOVSESPWLHAPSLKTSCEYVSLYHOPAAHMEDESDDDYN 115

Db 344 LGPAEAPLLNAPASASPQVLEAPWHPHPSLAKMCEYVSLGQFAPVNLDPDSDDDYN 403

Qy 116 VP 117

Db 404 IP 405

RESULT 5

AAE05349

ID AAG40735 standard; Protein; 407 AA.

XX

AC AAG40735;

XX

DT 18-OCT-2000 (first entry)

XX

DE Zea mays protein fragment SEQ ID NO: 50584.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence; corn.

XX

OS Zea mays subsp. mays.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123348.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136392.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.

KW acquired immune deficiency syndrome.

XX OS Homo sapiens.
XX PN W0200154733-AL.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001W0-US01312.
XX PR 31-JAN-2000; 2000US-0179065.
XX PR 04-FEB-2000; 2000US-0180628.
XX PR 24-FEB-2000; 2000US-0184684.
XX PR 02-MAR-2000; 2000US-0186350.
XX PR 16-MAR-2000; 2000US-0189874.
XX PR 17-MAR-2000; 2000US-0190076.
XX PR 18-APR-2000; 2000US-0198123.
XX PR 19-MAY-2000; 2000US-0205515.
XX PR 07-JUN-2000; 2000US-0209467.
XX PR 28-JUN-2000; 2000US-0214886.
XX PR 30-JUN-2000; 2000US-0215135.
XX PR 07-JUL-2000; 2000US-0216647.
XX PR 07-JUL-2000; 2000US-0216680.
XX PR 11-JUL-2000; 2000US-0217487.
XX PR 11-JUL-2000; 2000US-0217496.
XX PR 14-JUL-2000; 2000US-0218290.
XX PR 26-JUL-2000; 2000US-0220963.
XX PR 26-JUL-2000; 2000US-0220964.
XX PR 14-AUG-2000; 2000US-0224518.
XX PR 14-AUG-2000; 2000US-0224519.
XX PR 14-AUG-2000; 2000US-0225213.
XX PR 14-AUG-2000; 2000US-0225214.
XX PR 14-AUG-2000; 2000US-0225266.
XX PR 14-AUG-2000; 2000US-0225267.
XX PR 14-AUG-2000; 2000US-0225268.
XX PR 14-AUG-2000; 2000US-0225270.
XX PR 14-AUG-2000; 2000US-0225447.
XX PR 14-AUG-2000; 2000US-0225757.
XX PR 14-AUG-2000; 2000US-0225758.
XX PR 14-AUG-2000; 2000US-0225759.
XX PR 18-AUG-2000; 2000US-0226279.
XX PR 22-AUG-2000; 2000US-0226681.
XX PR 22-AUG-2000; 2000US-0226688.
XX PR 22-AUG-2000; 2000US-0227182.
XX PR 23-AUG-2000; 2000US-0227009.
XX PR 30-AUG-2000; 2000US-0228924.
XX PR 01-SEP-2000; 2000US-0229287.
XX PR 01-SEP-2000; 2000US-0229343.
XX PR 01-SEP-2000; 2000US-0229344.
XX PR 01-SEP-2000; 2000US-0229345.
XX PR 05-SEP-2000; 2000US-0229509.
XX PR 05-SEP-2000; 2000US-0229513.
XX PR 06-SEP-2000; 2000US-0230437.
XX PR 06-SEP-2000; 2000US-0230438.
XX PR 08-SEP-2000; 2000US-0231242.
XX PR 08-SEP-2000; 2000US-0231243.
XX PR 08-SEP-2000; 2000US-0231244.
XX PR 08-SEP-2000; 2000US-0231413.
XX PR 08-SEP-2000; 2000US-0231414.
XX PR 08-SEP-2000; 2000US-0232080.
XX PR 08-SEP-2000; 2000US-0232081.
XX PR 12-SEP-2000; 2000US-0231968.
XX PR 14-SEP-2000; 2000US-0232397.
XX PR 14-SEP-2000; 2000US-0232398.
XX PR 14-SEP-2000; 2000US-0232399.
XX PR 14-SEP-2000; 2000US-0232400.
XX PR 14-SEP-2000; 2000US-0232401.
XX PR 14-SEP-2000; 2000US-0233063.
XX PR 14-SEP-2000; 2000US-0233064.
XX PR 14-SEP-2000; 2000US-0233065.
XX PR 21-SEP-2000; 2000US-0234223.
XX PR 21-SEP-2000; 2000US-0234274.
XX PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 29-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2002, 17:28:11 : Search time 9.83333 Seconds
(without alignments)
586.214 Million cell updates/sec

Title: US-09-135-238B-2_COPY_18_253

Perfect score: 1254

Sequence: 1 RILPEVKVEGELGSGVTIKC.....HRQRALDYGSQSGREGQGF 236

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*

2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*

5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pep:*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1254	100.0	390	3	US-08-961-564A-2
2	459	36.6	107	3	US-08-961-564A-4
3	197	15.7	771	3	US-08-434-000A-8
4	197	15.7	771	4	US-09-312-157-8
5	187	14.9	109	3	US-08-961-564A-9
6	185	14.8	769	3	US-08-434-000A-10
7	185	14.8	769	4	US-09-312-157-10
8	180.5	14.4	608	4	US-09-095-385-4
9	180.5	14.4	746	3	US-08-434-000A-4
10	180.5	14.4	746	4	US-09-312-157-4
11	179.5	14.3	757	3	US-08-434-000A-6
12	179.5	14.3	757	4	US-09-312-157-6
13	166.5	13.3	624	2	US-08-642-406A-22
14	166.5	13.3	624	4	US-09-199-534-22
15	166.5	13.3	773	3	US-08-434-000A-2
16	166.5	13.3	773	4	US-09-312-157-2
17	96.5	7.7	476	3	US-08-487-550-4
18	90	7.2	453	3	US-08-466-151-8
19	90	7.2	453	4	US-08-466-163B-8
20	89.5	7.1	451	2	US-08-887-352B-14
21	89.5	7.1	451	2	US-08-887-352B-16
22	89.5	7.1	451	3	US-08-466-151-65
23	89.5	7.1	451	4	US-09-109-207C-14
24	89.5	7.1	451	4	US-09-109-207C-16
25	89.5	7.1	451	4	US-09-296-005-16
26	89.5	7.1	451	4	US-09-296-005-16
-	86.5	6.9	46	3	US-08-955-937A-10

Sequence 10, Appl
Sequence 20, Appl
Sequence 20, Appl
Sequence 20, Appl
Sequence 25, Appl
Sequence 25, Appl
Sequence 18, Appl
Patent No. 520236
Sequence 4, Appl
Sequence 4, Appl
Sequence 12, Appl
Sequence 21, Appl
Sequence 13, Appl
Sequence 26, Appl
Sequence 4, Appl
Sequence 26, Appl
Sequence 27, Appl

28 86.5 6.9 46 4 US-09-300-985-10
29 85.5 6.8 229 2 US-08-887-352B-20
30 85.5 6.8 229 4 US-09-109-207C-20
31 85.5 6.8 229 4 US-09-296-005-20
32 85.5 6.8 233 2 US-08-887-352B-25
33 85.5 6.8 233 4 US-09-109-207C-25
34 85.5 6.8 233 4 US-09-296-005-25
35 84.5 6.7 87 3 US-08-554-840-18
36 84.5 6.7 744 6 520236-25
37 84 6.7 195 3 US-08-955-937A-4
38 84 6.7 195 4 US-09-300-985-4
39 84 6.7 275 2 US-08-511-485-12
40 84 6.7 275 3 US-08-836-134-21
41 84 6.7 449 1 US-08-458-516-13
42 83.5 6.7 461 2 US-08-463-587A-26
43 83.5 6.7 461 2 US-08-463-667A-4
44 83.5 6.7 461 3 US-08-923-854-26
45 83.5 6.7 461 5 PCT-US91-09133-27

ALIGNMENTS

RESULT 1

US-08-961-564A-2
; Sequence 2, Application US/08961564A
; Patent No. 6114515
; GENERAL INFORMATION:
; APPLICANT: WU, SHUJIAN
; APPLICANT: SWEET, RAYMOND
; APPLICANT: TRUNER, ALEMGED
; TITLE OF INVENTION: PIGRL-1, A MEMBER OF IMMUNOGLOBULIN
; TITLE OF INVENTION: GENE SUPERFAMILY
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATHER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,564A
; FILING DATE: 30-OCT-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/056,935
; FILING DATE: 25-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70236
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 390 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-961-564A-2

Query Match 100.0%; Score 1254; DB 3; Length 390;
Best Local Similarity 100.0%; Pred. No. 1.le-119;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



QY 1 RILPEVKVEGELGSSVTIKCPLPEMHVRIYLCREMAGSGTCGTVSTTNFIKAEYKGRVT 60
Db 18 RILPEVKVEGELGSSVTIKCPLPEMHVRIYLCREMAGSGTCGTVSTTNFIKAEYKGRVT 77
QY 61 LKQYPRKNLFLVEVTQLTESDSGVYACGAGMNTDRGKTOKVTNLNVHSEYEPSWEEQPMPE 120
Db 78 LKQYPRKNLFLVEVTQLTESDSGVYACGAGMNTDRGKTOKVTNLNVHSEYEPSWEEQPMPE 137
QY 121 TPKWFHLPYLPOMPAYASSSKFVTRVTPAQRGKVPVHHSSPTTQITHRPRVSRASSVA 180
Db 138 TPKWFHLPYLPOMPAYASSSKFVTRVTPAQRGKVPVHHSSPTTQITHRPRVSRASSVA 197
QY 181 GDKPRTLPSTASKISALEGLKQPOTPSYNNHHTLRHQALDYGSOSGREGOGFH 236
Db 198 GDKPRTLPSTASKISALEGLKQPOTPSYNNHHTLRHQALDYGSOSGREGOGFH 253

RESULT 2

US-08-961-564A-4
; Sequence 4, Application US/08961564A
; Patent No. 6114515
; GENERAL INFORMATION:
; APPLICANT: WU, SHUJIAN
; APPLICANT: SWEET, RAYMOND
; APPLICANT: TRUNEH, ALEMSEGED
; TITLE OF INVENTION: FIGURE-1, A MEMBER OF IMMUNOGLOBULIN
; TITLE OF INVENTION: GENE SUPERFAMILY
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,564A
; FILING DATE: 30-OCT-1997
; CLASSIFICATION: 536

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/056,935
; FILING DATE: 25-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70236
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169

INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 107 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-961-564A-4

Query Match 36.6%; Score 459; DB 3; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.7e-39;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILPEVKVEGELGSSVTIKCPLPEMHVRIYLCREMAGSGTCGTVSTTNFIKAEYKGRVT 60
Db 18 RILPEVKVEGELGSSVTIKCPLPEMHVRIYLCREMAGSGTCGTVSTTNFIKAEYKGRVT 77

QY 61 LKQYPRKNLFLVEVTQLTESDSGVYACG 88
Db 78 LKQYPRKNLFLVEVTQLTESDSGVYACG 105

RESULT 3

US-08-434-000A-8
; Sequence 8, Application US/08434000A
; Patent No. 6046037
; GENERAL INFORMATION:
; APPLICANT: ANDREW C. HIATT, JULIAN
; APPLICANT: K.-C. MA, THOMAS LEHNER
; TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
; TITLE OF INVENTION: PROTEINS IN PLANTS AND THEIR USES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071

COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,000A
; FILING DATE:
; CLASSIFICATION: 435

PRIOR APPLICATION DATA:
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below:
; APPLICATION NUMBER: 08/367,395
; FILING DATE: 12/30/94
; ATTORNEY/AGENT INFORMATION:
; NAME: Guise, Jeffrey W.
; REGISTRATION NUMBER: 34,613
; REFERENCE/DOCKET NUMBER: 212/127
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 552-8400
; TELEFAX: (619) 552-0159
; TELEX: 67-3510

SEQUENCE LISTING
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 771 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; TOPOLOGY: DESCRIPTION:
; MOUSE Polyimmunoglobulin Receptor
US-08-434-000A-8

Query Match 15.7%; Score 197; DB 3; Length 771;
Best Local Similarity 32.4%; Pred. No. 1.3e-11;
Matches 48; Conservative 22; Mismatches 54; Indels 24; Gaps 5;

QY 13 GGSVTIKCPLPE----MHVRIYLCREMAGSGTCGTVSTTNFIKAEYKGRVTLKQYPRKN 68
Db 33 GDSVITCYYPDTSVNRHTRKYWCROGA-SGMCTTLISSNGYLSKEYSGRANLINFPENN 91
QY 69 LFLVEVTQLTESDSGVYACGAGMNTDRGKTOKVTNLNVHSEYEPSWEEQPMPTPKWHL 128
Db 92 TVVINIEQTQDTSYKCGLG-TSNRGLSFDVSLEV-----SOVPELPSDTHV- 139

QY 129 YLFQMPAYASSSKFVTRVTPAQRGKVP 156
Db 140 -----YTKDIGRNVITBCPFKRENV 160

1

2

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2002, 17:28:11 ; Search time 16.25 Seconds
(without alignments)
586.214 Million cell updates/sec

Title: US-09-135-238B-2
Perfect score: 2055
Sequence: 1 MDRWLWPLYFLVSGALRIL.....HQPAAMWEDSDDDYINVPA 390

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2047	99.6	390	3 US-08-961-564A-2	Sequence 2, Appli
2	550	26.8	107	3 US-08-961-564A-4	Sequence 4, Appli
3	198	9.6	771	3 US-08-434-000A-8	Sequence 8, Appli
4	198	9.6	771	4 US-09-312-157-8	Sequence 8, Appli
5	187	9.1	109	3 US-08-961-564A-9	Sequence 9, Appli
6	185	9.0	769	3 US-08-434-000A-10	Sequence 10, Appl
7	185	9.0	769	4 US-09-312-157-10	Sequence 10, Appl
8	182	8.9	608	4 US-09-095-385-4	Sequence 4, Appli
9	182	8.9	746	3 US-08-434-000A-4	Sequence 4, Appli
10	182	8.9	746	4 US-09-312-157-4	Sequence 4, Appli
11	182	8.9	757	3 US-08-434-000A-6	Sequence 6, Appli
12	182	8.9	757	4 US-09-312-157-6	Sequence 6, Appli
13	166.5	8.1	624	2 US-08-642-406A-22	Sequence 22, Appl
14	166.5	8.1	624	4 US-09-199-534-22	Sequence 22, Appl
15	166.5	8.1	773	3 US-08-434-000A-2	Sequence 2, Appli
16	166.5	8.1	773	4 US-09-312-157-2	Sequence 2, Appli
17	117.5	5.7	476	3 US-08-487-550-4	Sequence 4, Appli
18	117	5.7	453	3 US-08-466-151-8	Sequence 8, Appli
19	117	5.7	453	4 US-08-466-163B-8	Sequence 8, Appli
20	113.5	5.5	2337	3 US-08-713-118-2	Sequence 2, Appli
21	113.5	5.5	2337	4 US-09-452-007-2	Sequence 2, Appli
22	112	5.5	451	2 US-08-887-352B-14	Sequence 14, Appl
23	112	5.5	451	2 US-08-887-352B-16	Sequence 16, Appl
24	112	5.5	451	3 US-08-466-151-65	Sequence 65, Appl
25	112	5.5	451	4 US-09-109-207C-14	Sequence 14, Appl
26	112	5.5	451	4 US-09-109-207C-16	Sequence 16, Appl
27	112	5.5	451	4 US-09-296-005-14	Sequence 14, Appl

28	112	5.5	451	4 US-09-296-005-16	Sequence 16, Appl
29	111	5.4	476	2 US-08-378-939-10	Sequence 10, Appl
30	109.5	5.3	2339	1 US-08-455-543A-47	Sequence 47, Appl
31	109.5	5.3	2339	2 US-08-223-305C-47	Sequence 47, Appl
32	109.5	5.3	2339	4 US-09-268-163-6	Sequence 6, Appli
33	109.5	5.3	2343	4 US-09-268-163-4	Sequence 4, Appli
34	107	5.2	292	4 US-09-345-468-18	Sequence 18, Appl
35	107	5.2	313	4 US-09-345-468-16	Sequence 16, Appl
36	107	5.2	449	1 US-08-458-516-13	Sequence 13, Appl
37	106	5.2	390	2 US-08-979-424-1	Sequence 1, Appli
38	105	5.1	451	2 US-08-887-352B-18	Sequence 18, Appl
39	105	5.1	451	4 US-09-109-207C-18	Sequence 18, Appl
40	105	5.1	451	4 US-09-282-505-2	Sequence 2, Appli
41	105	5.1	451	4 US-09-054-255-2	Sequence 2, Appli
42	105	5.1	451	4 US-09-296-005-18	Sequence 18, Appl
43	104	5.1	554	1 US-08-347-254-1	Sequence 1, Appli
44	104	5.1	554	2 US-08-464-463-1	Sequence 1, Appli
45	103.5	5.0	459	1 US-08-157-101A-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-08-961-564A-2
; Sequence 2, Application US/08961564A
; Patent No. 6114515
; GENERAL INFORMATION:
; APPLICANT: WU, SHUJIAN
; APPLICANT: SWEET, RAYMOND
; APPLICANT: TRUNEH, ALEMSEGED
; TITLE OF INVENTION: PIGRL-1, A MEMBER OF IMMUNOGLOBULIN
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,564A
; FILING DATE: 30-OCT-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/056,935
; FILING DATE: 25-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70236
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 390 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-961-564A-2

Query Match 99.6%; Score 2047; DB 3; Length 390;
Best Local Similarity 99.7%; Pred. No. 2e-178;
Matches 389; Conservative 0; Mismatches 1; Indels 0; Gaps 0;



QY 1 MDRWLWPLFLPVSGALRILPEVKVEGELGGSVTKICPLPEMHVRIYLCREMAGSGTCGT 60
DB 1 MDRWLWPLFLPVSGALRILPEVKVEGELGGSVTKICPLPEMHVRIYLCREMAGSGTCGT 60
QY 61 VYSTNFIKAEYKGRVTLKQYPRKNLFLVEVTQJTESDGVYACGAGMNTDRGKTQKVT 120
DB 61 VYSTNFIKAEYKGRVTLKQYPRKNLFLVEVTQJTESDGVYACGAGMNTDRGKTQKVT 120
QY 121 NVHSEYFESWEEQPMPEPKFHLPLFQMPAYASSSKFVTRVTPAQRGKVPVPHSSP 180
DB 121 NVHSEYFESWEEQPMPEPKFHLPLFQMPAYASSSKFVTRVTPAQRGKVPVPHSSP 180
QY 181 TQIITHRVRVSRASSVAGDKPRTLPSTASKISALEGLLKPQTPSYNHHTRLHRQALD 240
DB 181 TQIITHRVRVSRASSVAGDKPRTLPSTASKISALEGLLKPQTPSYNHHTRLHRQALD 240
QY 241 YGSQSGREGQGFHILPTILGLFLALLGLVVKRAVERRKALSRARRLAVMRALSSQ 300
DB 241 YGSQSGREGQGFHILPTILGLFLALLGLVVKRAVERRKALSRARRLAVMRALSSQ 300
QY 301 PRGSPRPRSONNIYACPRRARGAAGTGEAPVPGGAPLPAPLOVSESPWLHAPSL 360
DB 301 PRGSPRPRSONNIYACPRRARGAAGTGEAPVPGGAPLPAPLOVSESPWLHAPSL 360
QY 361 KTSCEYVSLYHOPAAAMMEDSDDDYINVPA 390
DB 361 KTSCEYVSLYHOPAAAMMEDSDDDYINVPA 390

RESULT 2

US-08-961-564A-4
; Sequence 4, Application US/08961564A
; Patent No. 6114515

GENERAL INFORMATION:

APPLICANT: WU, SHUJIAN
APPLICANT: SWEET, RAYMOND
APPLICANT: TRUNEH, ALEMSEGED
TITLE OF INVENTION: PIGRL-1, A MEMBER OF IMMUNOGLOBULIN
TITLE OF INVENTION: GENE SUPERFAMILY
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: RATNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA: US/08/961,564A
FILING DATE: 30-OCT-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/056,935
FILING DATE: 25-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70236
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-961-564A-4

Query Match 26.8%; Score 550; DB 3; Length 107;
Best Local Similarity 99.0%; Pred. No. 6.6e-43;
Matches 104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDRWLWPLFLPVSGALRILPEVKVEGELGGSVTKICPLPEMHVRIYLCREMAGSGTCGT 60
DB 1 MDRWLWPLFLPVSGALRILPEVKVEGELGGSVTKICPLPEMHVRIYLCREMAGSGTCGT 60
QY 61 VYSTNFIKAEYKGRVTLKQYPRKNLFLVEVTQJTESDGVYACG 105
DB 61 VYSTNFIKAEYKGRVTLKQYPRKNLFLVEVTQJTESDGVYACG 105

RESULT 3

US-08-434-000A-8
; Sequence 8, Application US/08434000A
; Patent No. 6046037

GENERAL INFORMATION:

APPLICANT: ANDREW C. HIATT, JULIAN
APPLICANT: K.-C. MA, THOMAS LEHNER
TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
TITLE OF INVENTION: PROTEINS IN PLANTS AND THEIR USES
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/434,000A
FILING DATE:
CLASSIFICATION: 435

PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 1

GENERAL INFORMATION:

APPLICANT: ANDREW C. HIATT, JULIAN
APPLICANT: K.-C. MA, THOMAS LEHNER
TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
TITLE OF INVENTION: PROTEINS IN PLANTS AND THEIR USES
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071

Query Match 9.6%; Score 198; DB 3; Length 771;
Best Local Similarity 31.5%; Pred. No. 1.2e-09;
Matches 53; Conservative 23; Mismatches 66; Indels 24; Gaps 5;

QY 10 FLPVSGALRILPEVKVEGELGGSVTKICPLPE----MHVRIYLCREMAGSGTCGTVSTT 65

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2002, 17:24:41 ; Search time 21.2527 Seconds
(without alignments)
1921.018 Million cell updates/sec

Title: US-09-135-238b-2_COPY_18_253

Perfect score: 1254

Sequence: 1 RILPEVKVGEGLGSGVTIKC.....HQRALDYGSQSGREGQGFH 236

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPREMBL_19:*
- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1254	100.0	390	4	O60667
2	654	52.2	422	11	Q9D8T1
3	181	14.4	758	6	Q9NZH7
4	175.5	14.0	534	4	Q96SA2
5	166	13.2	455	11	Q920L8
6	166	13.2	535	11	Q9EQT7
7	121	9.6	299	4	Q9UBK4
8	121	9.6	299	4	Q9UGN4
9	117.5	9.4	359	4	O43656
10	111.5	8.9	307	11	Q94947
11	108	8.6	364	4	Q96D42
12	106	8.5	298	4	Q9HD97
13	106	8.5	301	4	Q95100
14	105	8.4	270	4	Q9DWT1
15	105	8.4	270	4	Q9H564
16	100.5	8.0	503	4	Q9H8V0

17	100.5	8.0	504	4	Q9UD78	Q9ud78 homo sapien
18	100	8.0	227	11	Q9JL34	Q9jl34 mus musculu
19	98.5	7.9	230	4	Q9NZC2	Q9nzc2 homo sapien
20	98	7.8	886	13	P70049	P70049 xenopus lae
21	97.5	7.8	662	4	O60926	O60926 homo sapien
22	96	7.7	230	11	Q9JKE2	Q9jke2 mus musculu
23	96	7.7	589	4	Q9H802	Q9h802 homo sapien
24	95	7.6	258	4	Q9DWT2	Q9dwt2 homo sapien
25	95	7.6	258	4	Q9H563	Q9h563 homo sapien
26	94	7.5	227	11	Q9NNH8	Q9nnh8 mus musculu
27	93.5	7.5	1847	5	Q9NKN5	Q9nkn5 leishmania
28	92.5	7.4	336	13	Q90Z89	Q90z89 brachydanio
29	92.5	7.4	504	11	O60786	O60786 mus musculu
30	92	7.3	287	11	Q9D7B8	Q9d7b8 mus musculu
31	91.5	7.3	502	4	Q9UD77	Q9ud77 homo sapien
32	91.5	7.3	502	4	Q12800	Q12800 homo sapien
33	91.5	7.3	4072	5	Q9W4Y4	Q9w4y4 drosophila
34	91	7.3	227	11	Q9NNH9	Q9nnh9 mus musculu
35	91	7.3	757	5	Q9V6T1	Q9v6t1 drosophila
36	90.5	7.2	315	4	O96K52	O96k52 homo sapien
37	90.5	7.2	549	5	Q9U3F1	Q9u3f1 caenorhabdi
38	90.5	7.2	553	5	Q9GTC3	Q9gtc3 caenorhabdi
39	90.5	7.2	556	5	Q9XXU7	Q9xxu7 caenorhabdi
40	90.5	7.2	614	4	O96FE5	O96fe5 homo sapien
41	90.5	7.2	614	11	Q9D1T0	Q9d1t0 mus musculu
42	90.5	7.2	754	4	O96KV7	O96kv7 homo sapien
43	90.5	7.2	1209	11	Q924W6	Q924w6 mus musculu
44	90	7.2	822	5	O61674	O61674 drosophila
45	89.5	7.1	1417	12	O67631	O67631 gallid herp

ALIGNMENTS

RESULT 1

O60667 ID O60667 PRELIMINARY; PRT; 390 AA.
AC O60667;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ANTI-FAS-INDUCED APOPTOSIS (REGULATOR OF FAS-INDUCED APOPTOSIS).
GN TOSO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98245048; PubMed=9586636;
RA Hitoshi Y., Lorens J., Kitada S.I., Fisher J., Labarge M., Ring H.Z.,
RA Francke U., Reed J.C., Kinoshita S., Nolan G.P.;
RT "Toso, a cell surface, specific regulator of Fas-induced apoptosis in
T cells";
RL Immunity 8:461-471(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=PRIMARY B-CELLS FROM TONSILS;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF057557; AAC18830.1; -;
DR EMBL; BC006401; AAH06401.1; -;
DR InterPro; IPR003599; Ig_
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00409; Ig; 1.
SQ SEQUENCE 390 AA; 43146 MW; F891D217EECA99C6 CRC64;

Query Match 100.0%; Score 1254; DB 4; Length 390;

Best Local Similarity 100.0%; Pred. No. 1.7e-115;

Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RILPEVKVGEGLGSGVTIKCPLPMHVRIVLCREMAGSGTCGVTVVSTNFIKAEYKGRVT 60

```

Db 18 RILPEVKVELEGSGVTIKCPPEMHVRIYLCREMAGSGTCGTVTTFIKAEYKGRVT 77
QY 61 LKQYPRKNLFLVEVTLTSDSGVYACGAGMNTDRGKTQKVTNLNVHSEYEPWESQPMPE 120
Db 78 LKQYPRKNLFLVEVTLTSDSGVYACGAGMNTDRGKTQKVTNLNVHSEYEPWESQPMPE 137
QY 121 TPKEFHLPYLFPQMPAYASSKFVTRVTPAQRGKVPVPHHSSPTQITHRPRVRSASSVA 180
Db 138 TPKEFHLPYLFPQMPAYASSKFVTRVTPAQRGKVPVPHHSSPTQITHRPRVRSASSVA 197
QY 181 GDKPRTFLPSTTASIKALEGLLKQPOTPSYNNHTRLHRORALDYGSGSGREGQGFH 236
Db 198 GDKPRTFLPSTTASIKALEGLLKQPOTPSYNNHTRLHRORALDYGSGSGREGQGFH 253

RESULT 2
Q9DBT1 PRELIMINARY; PRT; 422 AA.
AC Q9DBT1 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE 1810037B05RIK PROTEIN.
GN 1810037B05RIK
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 1
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull D., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Sakamoto N.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.
RA "Functional annotation of a full-length mouse cDNA collection.";
RT Nature 403:685-690(2001).
DR EMBL: AK007714; BAB25207.1; -.
DR MGD: MGI:1916419; 1810037B05RIK.
DR InterPro: IPR003600; Ig_like.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: Pf00047; Ig; 1.
DR SMART: SM00410; Ig_like; 1.
SQ SEQUENCE 422 AA; 47532 MW; 2597083A50AD8B6E CRC64;

Query Match 52.2%; Score 654; DB 11; Length 422;
Best Local Similarity 53.3%; Pred. No. 3.2e-56;
Matches 129; Conservative 34; Mismatches 71; Indels 8; Gaps 4;

```

```

QY 1 RILPEVKVELEGSGVTIKCPPEMHVRIYLCREMAGSGTCGTVTTFIKAEYKGRVT 60
Db 18 RVLPEVLNVEGSGIIIECPQLQHLHVMYLCRQMAKPGICSTVVSNT-FVKKEYERRV 76
QY 61 LKQYPRKNLFLVEVTLTSDSGVYACGAGMNTDRGKTQKVTNLNVHSEY-EPWESQPMPE 119
Db 77 LTPCLDKKFLVEMTQLTENDDGIYACGVGMKTKDGKTQKITLNVHNEYPEPPEWDEWTS 136

```

```

QY 120 ETPKWEHLPYLFQMP-----AYASSKFVTRVTPAQRGKVPVPHHSSPTQITHRPRVS 174
Db 137 ERPRHLRFLQHPWLGSHSEHSSGVIAKYTTPTPKTEAPPVHPQSSITSVTQHPRVY 196
QY 175 RASVAGDKPRTFLPSTTASIKALEGLLKQPOTPSYNNHTRLHRORALDYGSGSGREGQ 234
Db 197 RAFSVSATKSPALLPATPTASKTSTQQA-IRPLEASYSHTLRLHQRHGHGPHYGRDGR 255
QY 235 FH 236
Db 256 LH 257

RESULT 3
Q9N2H7 PRELIMINARY; PRT; 758 AA.
AC Q9N2H7 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE POLY-IG RECEPTOR PRECURSOR.
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN 1
SEQUENCE FROM N.A.
RA Sone T., Kumura H.;
RT "Porcine mammary gland cDNA clone, similar to poly-Ig receptor.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB032195; BAA84283.2; -.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003600; Ig_like.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: Pf00047; Ig; 5.
DR SMART: SM00409; IG; 3.
DR SMART: SM00410; Ig_like; 2.
KW Signal; Receptor.
FT SIGNAL 1 18
FT SIGNAL 18
SQ SEQUENCE 758 AA; 83154 MW; D5BEBIA8B082D247 CRC64;

Query Match 14.4%; Score 181; DB 6; Length 758;
Best Local Similarity 33.9%; Pred. No. 3.4e-09;
Matches 40; Conservative 22; Mismatches 40; Indels 16; Gaps 4;

```

```

QY 2 ILPEVKVEGL-----GGSVTIKCP-----EMHVRIYLCREMAGSGTCGTVYST 47
Db 12 IFPVVSMKSPIFQPDVSSVEGSSVIRCYIPATSVNHRHRSKYWCR-IGAKGRCTLISS 70
QY 48 TNFIKAEYKGRVTLKQYPRKNLFLVEVTLTSDSGVYACGAGMNTDRGKTQKVTN 105
Db 71 EGYISKDYKGRANLTFNPENGTFVMDIGHTRGDSGLYKGLGISS-RGLSPDVSLEV 127

RESULT 4
Q96SA2 PRELIMINARY; PRT; 534 AA.
AC Q96SA2 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE FKSG87 PROTEIN.
GN FKSG87.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 1
SEQUENCE FROM N.A.
RA Wang Y.-G., Gong L.;
RT "Molecular cloning and characterization of FKSG87, a novel gene
located on human chromosome 1.";
Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RL

```



```
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00410; IG_Like; 1.
SQ SEQUENCE 270 AA; 29668 MW; 32791A7B49367E46 CRC64;

Query Match      8.4%; Score 105; DB 4; Length 270;
Best Local Similarity 21.8%; Pred. No. 0.028;
Matches 58; Conservative 31; Mismatches 97; Indels 80; Gaps 11;

Qy 13 GGSVTIKCPLP---EMHVRIYLCREMAGSGTCGTVVSTTNFIKAEYKGRVTLKQYPRKNL 69
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 33 GQTLTVRCQYPTGSLYEKKGWCKE-ASALVCIRLVTSKSPRTMAWTSRETINDDPDAGF 91
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 70 FLVEVTQLTESDSGVYACGAGMNTDRGKTQKVTNLNVHSEYEPSEWEEQPMPETPKWFHLPY 129
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 92 FTVTMTDLREEDSGHYWCRIYRPSDINSVSKSVRF-----Y 126
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 130 LFQMPAYAS-----SSKEVTVRTVTPAQRG-----KVP-PVHHSSP-TTQI 167
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 127 LVVSPASASTQTSWTPRDLVSSQTQTOSCVPTAGARQAPESPSTIPVPSHPSSPLDVPL 186
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 168 THRPRVS--RASSVAGDKPRTFLP---STTASKISALEGLL----- 203
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 187 PSRPQNSTLRPGPAA---PIALVPVFCGLLVAKSLVLSALLVWVLRNRHMQHGRSLH 243
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 204 --KPQTPSYNHHTRLHRQRALDYGQ 227
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 244 PAQPRQAHRFPLSHRAPGGTYGK 269
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

Search completed: October 28, 2002, 17:32:36
Job time : 24.2527 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2002, 17:28:11 ; Search time 9.83333 Seconds
(without alignments)
586.214 Million cell updates/sec

Title: us-09-135-238b-2_COPY_18_253

Perfect score: 1254

Sequence: 1 RLPEVKVEGLGGSVTKK.....HRQRALDYGSOSGRGQGFH 236

Scoring table: BLOSUM62

Gapop 10.0 , Capext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2.6/ptodata/1/iaa/5A_COMB.pep.*

2: /cgn2.6/ptodata/1/iaa/5B_COMB.pep.*

3: /cgn2.6/ptodata/1/iaa/6A_COMB.pep.*

4: /cgn2.6/ptodata/1/iaa/6B_COMB.pep.*

5: /cgn2.6/ptodata/1/iaa/PCUS_COMB.pep.*

6: /cgn2.6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1254	100.0	350	US-08-961-564A-2	Sequence 2, Appli
2	459	36.6	107	US-08-961-564A-4	Sequence 4, Appli
3	197	15.7	771	US-08-434-000A-8	Sequence 8, Appli
4	197	15.7	771	US-09-312-157-8	Sequence 8, Appli
5	187	14.9	109	US-08-961-564A-9	Sequence 9, Appli
6	185	14.8	769	US-08-434-000A-10	Sequence 10, Appl
7	185	14.8	769	US-09-312-157-10	Sequence 10, Appl
8	180.5	14.4	608	US-09-095-385-4	Sequence 4, Appli
9	180.5	14.4	746	US-08-434-000A-4	Sequence 4, Appli
10	180.5	14.4	746	US-09-312-157-4	Sequence 4, Appli
11	179.5	14.3	757	US-08-434-000A-6	Sequence 6, Appli
12	179.5	14.3	757	US-09-312-157-6	Sequence 6, Appli
13	166.5	13.3	624	US-08-642-406A-22	Sequence 22, Appl
14	166.5	13.3	624	US-09-199-534-22	Sequence 22, Appl
15	166.5	13.3	773	US-08-434-000A-2	Sequence 2, Appli
16	166.5	13.3	773	US-09-312-157-2	Sequence 2, Appli
17	96.5	7.7	476	US-08-487-550-4	Sequence 4, Appli
18	90	7.2	453	US-08-466-151-8	Sequence 8, Appli
19	90	7.2	453	US-08-466-163B-8	Sequence 8, Appli
20	89.5	7.1	451	US-08-887-352B-14	Sequence 14, Appl
21	89.5	7.1	451	US-08-887-352B-16	Sequence 16, Appl
22	89.5	7.1	451	US-08-466-151-65	Sequence 65, Appl
23	89.5	7.1	451	US-09-109-207C-14	Sequence 14, Appl
24	89.5	7.1	451	US-09-109-207C-16	Sequence 16, Appl
25	89.5	7.1	451	US-09-296-005-14	Sequence 14, Appl
26	89.5	7.1	451	US-09-296-005-16	Sequence 16, Appl
27	86.5	6.9	46	US-08-955-937A-10	Sequence 10, Appl

28 86.5 6.9 46 4 US-09-300-985-10 Sequence 10, Appli
29 85.5 6.8 229 2 US-08-887-352B-20 Sequence 20, Appl
30 85.5 6.8 229 4 US-09-109-207C-20 Sequence 20, Appl
31 85.5 6.8 229 4 US-09-296-005-20 Sequence 20, Appl
32 85.5 6.8 233 2 US-08-887-352B-25 Sequence 25, Appl
33 85.5 6.8 233 4 US-09-109-207C-25 Sequence 25, Appl
34 85.5 6.8 233 4 US-09-296-005-25 Sequence 25, Appl
35 84.5 6.7 87 3 US-08-554-840-18 Patent No. 5202236
36 84.5 6.7 744 6 5202236-25
37 84 6.7 195 3 US-08-955-937A-4 Sequence 4, Appli
38 84 6.7 195 4 US-09-300-985-4 Sequence 4, Appli
39 84 6.7 275 2 US-08-511-485-12 Sequence 12, Appl
40 84 6.7 275 3 US-08-836-134-21 Sequence 21, Appl
41 84 6.7 449 1 US-08-458-516-13 Sequence 13, Appl
42 83.5 6.7 461 2 US-08-463-587A-26 Sequence 26, Appl
43 83.5 6.7 461 2 US-08-923-854-26 Sequence 26, Appl
44 83.5 6.7 461 3 US-08-923-854-26 Sequence 26, Appl
45 83.5 6.7 461 5 PCT-US91-09133-27 Sequence 27, Appl

ALIGNMENTS

RESULT 1

US-08-961-564A-2
; Sequence 2, Application US/08961564A
; Patent No. 6114515
; GENERAL INFORMATION:
; APPLICANT: WU, SHUJIAN
; APPLICANT: SWEET, RAYMOND
; APPLICANT: TRUNEH, ALEMGED
; TITLE OF INVENTION: PIGRL-1, A MEMBER OF IMMUNOGLOBULIN
; TITLE OF INVENTION: GENE SUPERFAMILY
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,564A
; FILING DATE: 30-OCT-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/056,935
; FILING DATE: 25-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70236
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 390 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-961-564A-2

Query Match 100.0%; Score 1254; DB 3; Length 390;
Best Local Similarity 100.0%; Pred. No. 1.1e-119;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,000A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below: 1
; APPLICATION NUMBER: 08/367,395
; FILING DATE: 12/30/94
; ATTORNEY/AGENT INFORMATION:
; NAME: Guise, Jeffrey W.
; REGISTRATION NUMBER: 34,613
; REFERENCE/DOCKET NUMBER: 212/127
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 552-8400
; TELEFAX: (619) 552-0159
; TELEX: 67-3510
; SEQUENCE LISTING
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 769 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; DESCRIPTION: Rat Polyimmunoglobulin Receptor
; US-08-434-000A-10

Query Match 14.8%; Score 185; DB 3; Length 769;
Best Local Similarity 34.5%; Pred. No. 2.2e-10;
Matches 41; Conservative 21; Mismatches 41; Indels 16; Gaps 4;

QY 13 GGSVTIKPLPE----MHVRIYLCREMAGSGTCGTIVSTTNEIKAEYKGRVTLKQYPRKN 68
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 33 GNSVITCYPTDSVNRHTRKYWCROGA-NGYCATLISSNGYLSKEYSGRASLINFPENS 91
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 69 LFLVEVTLTSDSGVYACGAGMNTDRGKTQKTVLNVHSEYEPSPWEEQPMPTPKWFHL 127
| : | : | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 92 TFVINIAHLTQEDTSGYKCGLG-TTNRGLFFDYSLEV-----SQVPEPPNDTHV 139

RESULT 7
US-09-312-157-10
; Sequence 10, Application US/09312157
; Patent No. 6303341
; GENERAL INFORMATION:
; APPLICANT: ANDREW C. HIATT, JULIAN
; K.-C. MA, THOMAS LEHNER
; TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
; PROTEINS IN PLANTS AND THEIR USES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/312,157
; FILING DATE: 14-May-1999
; CLASSIFICATION: <Unknown>

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/434,000
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Guise, Jeffrey W.
; REGISTRATION NUMBER: 34,613
; REFERENCE/DOCKET NUMBER: 212/127
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 552-8400
; TELEFAX: (619) 552-0159
; TELEX: 67-351
; SEQUENCE LISTING
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 769 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; DESCRIPTION: Rat Polyimmunoglobulin Receptor
; US-09-312-157-10

Query Match 14.8%; Score 185; DB 4; Length 769;
Best Local Similarity 34.5%; Pred. No. 2.2e-10;
Matches 41; Conservative 21; Mismatches 41; Indels 16; Gaps 4;

QY 13 GGSVTIKPLPE----MHVRIYLCREMAGSGTCGTIVSTTNEIKAEYKGRVTLKQYPRKN 68
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 33 GNSVITCYPTDSVNRHTRKYWCROGA-NGYCATLISSNGYLSKEYSGRASLINFPENS 91
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 69 LFLVEVTLTSDSGVYACGAGMNTDRGKTQKTVLNVHSEYEPSPWEEQPMPTPKWFHL 127
| : | : | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 92 TFVINIAHLTQEDTSGYKCGLG-TTNRGLFFDYSLEV-----SQVPEPPNDTHV 139

RESULT 8
US-09-095-385-4
; Sequence 4, Application US/09095385
; Patent No. 6300104
; GENERAL INFORMATION:
; APPLICANT: Morrison, Sherie L.
; APPLICANT: Chantalacharuvu, Kote R.
; TITLE OF INVENTION: SECRETORY IMMUNOGLOBULIN PRODUCED
; BY SINGLE CELLS AND METHODS FOR MAKING AND USING
; TITLE OF INVENTION: SAME
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
; STREET: 11150 Santa Monica Boulevard, Suite 400
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/095,385
; FILING DATE: 09-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/050,969
; FILING DATE: 19-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Canady, Karen S
; REGISTRATION NUMBER: 39,927
; REFERENCE/DOCKET NUMBER: 30435.45USU1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 310 445-1140
; TELEFAX: 310 445-9031
; TELEX:

```


Db 65 NLTFNPENGTFFVYVNIQAQLSDSGRYKCGLGINS-RGLSFDVSLEV 109

RESULT 11

US-08-434-000A-6

Sequence 6, Application US/08434000A

Patent No. 6046037

GENERAL INFORMATION:

APPLICANT: ANDREW C. HIATT, JULIAN

APPLICANT: K.-C. MA, THOMAS LEHNER

TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION

TITLE OF INVENTION: PROTEINS IN PLANTS AND THEIR USES

NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street

SUITE: Suite 4700

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/434,000A

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

PRIOR APPLICATION DATA: including application

PRIOR APPLICATION DATA: described below:

APPLICATION NUMBER: 08/367,395

FILING DATE: 12/30/94

ATTORNEY/AGENT INFORMATION:

NAME: Guise, Jeffrey W.

REGISTRATION NUMBER: 34,613

REFERENCE/DOCKET NUMBER: 212/127

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 552-8400

TELEFAX: (619) 552-0159

TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 757 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

DESCRIPTION: Bovine Polyimmunoglobulin Receptor

US-08-434-000A-6

Query Match 14.3%; Score 179.5; DB 3; Length 757;

Best Local Similarity 37.6%; Pred. No. 7.7e-10;

Matches 41; Conservative 19; Mismatches 42; Indels 7; Gaps 4;

QY 4 PEVKVEGELGGSVTIKCPPLP-----EMHVRIYLCREMAGSGTCGTGVVSTNFTKAEYKGRV 59

Db 25 PE-EVSSVEGRSVSIKCYPPPTSVNRHTRKYWCROGA-QGRCTTLISSEGVSDDYVGRA 82

QY 60 TLKQYPRKNLFVETQLTESDSGVYACGAGMNTDRGKTOKVTNLVHSE 108

Db 83 NLTFNPESGTFVVDISHLTHKDSGRYKCGLGISS-RGLNFDVSLEVSQD 130

RESULT 12

US-09-312-157-6

Sequence 6, Application US/09312157

Patent No. 6303341

GENERAL INFORMATION:

APPLICANT: ANDREW C. HIATT, JULIAN

K.-C. MA, THOMAS LEHNER

TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION

TITLE OF INVENTION: PROTEINS IN PLANTS AND THEIR USES

NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street

SUITE: Suite 4700

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/312,157

FILING DATE: 14-May-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/434,000

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Guise, Jeffrey W.

REGISTRATION NUMBER: 34,613

REFERENCE/DOCKET NUMBER: 212/127

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 552-8400

TELEFAX: (619) 552-0159

TELEX: 67-351

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 757 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

DESCRIPTION: Bovine Polyimmunoglobulin Receptor

US-09-312-157-6

Query Match 14.3%; Score 179.5; DB 4; Length 757;

Best Local Similarity 37.6%; Pred. No. 7.7e-10;

Matches 41; Conservative 19; Mismatches 42; Indels 7; Gaps 4;

QY 4 PEVKVEGELGGSVTIKCPPLP-----EMHVRIYLCREMAGSGTCGTGVVSTNFTKAEYKGRV 59

Db 25 PE-EVSSVEGRSVSIKCYPPPTSVNRHTRKYWCROGA-QGRCTTLISSEGVSDDYVGRA 82

QY 60 TLKQYPRKNLFVETQLTESDSGVYACGAGMNTDRGKTOKVTNLVHSE 108

Db 83 NLTFNPESGTFVVDISHLTHKDSGRYKCGLGISS-RGLNFDVSLEVSQD 130

RESULT 13

US-08-642-406A-22

Sequence 22, Application US/08642406A

Patent No. 5959177

GENERAL INFORMATION:

APPLICANT: Hein, Mich B.

APPLICANT: Hiatt, Andrew C.

APPLICANT: Ma, Julian K.C.

TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING ASSEMBLED

TITLE OF INVENTION: SECRETORY ANTIBODIES

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE

STREET: 10666 No. 5959177th Torrey Pines Road, Tpc-8

CITY: La Jolla

STATE: California


```
;
; TELEFAX: (619) 552-0159
; TELEX: 67-3510
; TELEX: SEQUENCE LISTING
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 773 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; DESCRIPTION: Rabbit polyimmunoglobulin receptor
US-08-434-000A-2

Query Match      13.3%; Score 166.5; DB 3; Length 773;
Best Local Similarity 42.7%; Pred. No. 1.7e-08;
Matches 44; Conservative 11; Mismatches 41; Indels 7; Gaps 4;

Qy 13 GGSVTIKCLP-...EMHVRIYLCHEMAGSGTGVVSTTFIKAEYKGRVTLKQYPRKN 68
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 39 GDSVSITCYPTTSVTRHSRKFECRE-EESGRCVTLAS-TGYTSQEYSGRGLTDFPDKG 96
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 69 LELVEVTOLTESDSGVYACGAGMNTDRGKTQKVTNLNVHSEYEP 111
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 97 EFVTVDTQNDSGSYKCGGVN-GRGLDFGVNLVSKPEP 138
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: October 28, 2002, 17:33:13
Job time : 12.8333 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2002, 17:24:41 ; Search time 10.6263 seconds
(without alignments)
1921.018 Million cell updates/sec

Title: US-09-135-238b-2_COPY_273_390
Perfect score: 616
Sequence: 1 KRAVERKALSRARRLAVR.....HOPAAEMDSDDDDYINVPA 118

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriaph:*
- 17: sp_bacteriap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	616	100.0	390	4	O60667
2	361.5	58.7	422	11	Q9D8T1
3	96	15.6	114	11	Q64561
4	96	15.6	115	11	Q64390
5	93	15.1	559	5	O46132
6	89.5	14.5	163	12	O9WSV8
7	89.5	14.5	748	10	O9AH83
8	88.5	14.4	526	16	O9RUB1
9	87.5	14.2	2936	5	Q9NKP7
10	87	14.1	821	4	O96JH3
11	85.5	13.9	166	2	O30801
12	85	13.8	523	5	O77238
13	85	13.8	554	5	O9VTX7
14	84.5	13.7	152	12	Q9JH32
15	84.5	13.7	219	4	O96PJ4
16	84.5	13.7	251	2	O9EY22

17	84.5	13.7	359	4	Q9BR57
18	84.5	13.7	1049	10	Q9ZNR9
19	84.5	13.7	2129	12	Q9J6K9
20	84	13.6	384	11	Q9CUF6
21	83	13.5	161	2	Q9F180
22	83	13.5	706	5	Q9VYR5
23	83	13.5	905	11	Q9WV47
24	83	13.5	1039	2	Q93L97
25	83	13.5	1260	5	Q9V468
26	83	13.5	1740	11	Q9JLU4
27	83	13.5	1806	11	Q9WUY7
28	82.5	13.4	225	4	Q96PU1
29	82.5	13.4	270	4	Q96PJ2
30	82.5	13.4	2115	12	Q86500
31	82	13.3	311	16	Q92RT2
32	82	13.3	386	4	Q9NQU0
33	82	13.3	392	4	Q75526
34	82	13.3	452	4	Q96SP6
35	82	13.3	802	5	Q9VJ74
36	81.5	13.2	163	12	Q9JG79
37	81.5	13.2	528	16	Q9RSJ1
38	81.5	13.2	2116	12	O40955
39	81.5	13.2	2301	10	Q9ATK5
40	81	13.1	148	4	Q96BZ7
41	81	13.1	394	2	Q9L0B0
42	81	13.1	465	4	Q9BWD7
43	81	13.1	513	4	Q96H99
44	81	13.1	553	16	Q9RSN4
45	81	13.1	772	2	Q9AD03

ALIGNMENTS

RESULT 1

O60667
ID O60667 PRELIMINARY; PRT; 390 AA.
AC O60667;
DC 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ANTI-FAS-INDUCED APOPTOSIS (REGULATOR OF FAS-INDUCED APOPTOSIS).
GN TOSO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98246048; PubMed=9586636;
RA Hitoshi Y., Lorens J., Kitada S.I., Fisher J., Labarge M., Ring H.Z.,
RA Francke U., Reed J.C., Kinoshita S., Nolan G.P.;
RT "Toso, a cell surface, specific regulator of Fas-induced apoptosis in
T cells.";
RL Immunity 8:461-471(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=PRIMARY B-CELLS FROM TONSILS;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF057557; AAC18830.1; -;
DR EMBL; BC006401; AAH06401.1; -;
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; ig; 1.
DR SMART; SM00409; IG; 1.
SQ SEQUENCE 390 AA; 43146 MW; FE91D217E9CA99C6 CRC64;

Query Match 100.0%; Score 616; DB 4; Length 390;
Best Local Similarity 100.0%; Pred. No. 7.3e-54;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KRAVERKALSRARRLAVRMRALRESSQPRGRSPPRRSONNIYSACPRRARGADAAGTGE 60


```
CC -I- SIMILARITY: STRONG, TO OTHER MAMMALIAN SPERMATID NUCLEAR
CC TRANSITION PROTEINS 2.
DR EMBL: 246939; CAA87064.1; -.
DR InterPro: IPR000678; TP2.
DR Pfam: PF01254; TP2; 1.
DR PROSITE: PS00970; TP2_1; 1.
DR PROSITE: PS00971; TP2_2; 1.
KW Chromosomal protein; DNA-binding; Metal-binding; Nuclear protein;
KW Nucleosome core; Spermatogenesis; Zinc.
SQ SEQUENCE 115 AA; 12948 MW; 29E180FF27894C1A CRC64;

Query Match 15.6%; Score 96; DB 11; Length 115;
Best Local Similarity 32.5%; Pred. No. 0.024; 36; Indels 6; Gaps 3;
Matches 25; Conservative 10; Mismatches 10;

QY 20 RMALESSO-RPRGSPRSONNIYACPRRARGADAAG---TGEAPVPGCAPLPAPL 75
   : : : : : | | | | | | | | | | | | | | | | | | | | | | | |
DB 4 KMSLTTPHPHSSSRPOSHTNNQCACSHHCSCSQAGHPSSSSPSGPTKPKPTM 63
   : : : : : | | | | | | | | | | | | | | | | | | | | | | | |
QY 76 QVSESPWLHAPSLKTS 92
   || || | | | | | | | | | | | | | | | | | | | | | | | |
DB 64 HSRYSR--SRPSHRGSC 78

RESULT 5
O46132 PRELIMINARY; PRT; 559 AA.
ID O46132
AC O46132;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NICOTINIC ACETYLCHOLINE RECEPTOR, ALPHA1 SUBUNIT (FRAGMENT).
GN NACHR.
OS Locusta migratoria (Migratory locust).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
OC Acridomorpha; Acridoidea; Acrididae; Locusta.
OX NCBI_TaxID=7004;
RN [1]
RP MEDLINE=98325051; PubMed=9660807;
RA Hermen B., Stetzer E., Thees R., Heiermann R., Schrattenholz A.,
RA Ebblinghaus U., Kretschmer A., Methfessel C., Reinhardt S.,
RA Maelicke A.;
RT "Neuronal nicotinic receptors in the locust locusta migratoria.";
RL J. Biol. Chem. 273:18394-18404(1998).
CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -I- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
DR EMBL: AJ000390; CAA04052.1; -.
DR InterPro: IPR000188; GABAA_receptor.
DR InterPro: IPR001175; Neur_chan.
DR Pfam: PF02931; Neur_chan_LBD; 1.
DR Pfam: PF02932; Neur_chan_memb; 1.
DR PRINTS: PR00252; NRIONCHANNEL.
DR PROSITE: PS00236; NEUOTR_ION_CHANNEL; 1.
KW Glycoprotein; ionic channel; Postsynaptic membrane; Receptor;
KW Transmembrane.
FT NON_TER 1
SQ SEQUENCE 559 AA; 62442 MW; 6B9B33F2778B3DE CRC64;

Query Match 15.1%; Score 93; DB 5; Length 559;
Best Local Similarity 26.9%; Pred. No. 0.26;
Matches 36; Conservative 12; Mismatches 60; Indels 26; Gaps 3;

QY 1 KRAVERKALSRARRLAVRMRALES-----SORPGSPRPRSONNIYACPR 48
   : | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 361 ERGLRRRPGRRGRAGAPQGLLQRPAAHQLQGRPLARQAPARRAAGQLPGARAAG 420
   : | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 49 RARGADAAGT-----EAPVPGCAPLPAPLQVSESPWLHA---PSLKTSCEY 94
   : | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 421 GAGGGAAYGRGGVQEPATAAAATASGPGVAPAGVRVSRPPAFPHSRCPPEVHRSCFC 480
   : | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 95 VSLYHQPAAMMEDS 108
```

```
Db 481 VRFTAEHTRMLEDS 494
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | |
RESULT 6
Q9WSV8 PRELIMINARY; PRT; 163 AA.
ID Q9WSV8
AC Q9WSV8;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE DNA, COMPLETE GENOME, ISOLATE:TTV SANBAN.
OS "Tt virus".
OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
OX NCBI_TaxID=68887;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TTV SANBAN;
RA Hijikata M.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TTV SANBAN;
RX MEDLINE=99335592; PubMed=10405352;
RA Hijikata M., Takahashi K., Mishiro S.;
RT "Complete circular DNA genome of a Tt virus variant (isolate name
RT SANBAN) and 44 partial ORF2 sequences implicating a great degree of
RT diversity beyond genotypes.";
RL Virology 260:17-22(1999).
DR EMBL: AB025946; BAA82148.1; -.
DR InterPro: IPR004118; TT_ORF2.
DR Pfam: PF02957; TT_ORF2; 1.
SQ SEQUENCE 163 AA; 17179 MW; 794454FBD979B17E CRC64;

Query Match 14.5%; Score 89.5; DB 12; Length 163;
Best Local Similarity 31.1%; Pred. No. 0.15;
Matches 32; Conservative 5; Mismatches 39; Indels 27; Gaps 4;

QY 5 ERKALSRARRLAVRMRALESQSRPGSPRPRSONNIYACPRRARGADAAG----- 57
   : | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 21 KKKKLLRLSLRAPQAPRRAM--SWRPVHDAPCIERNWYEAACFRAHAGSCGNGFIAHN 78
   : | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 58 -----TGEAPVPG---PGAP-----LPPAPLQVSESPW 82
   : | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 79 LLAGRYGFTGGPPPGPPGPPGTPQVQRASNSPAAPQPPALPW 121
   : | | | | | | | | | | | | | | | | | | | | | | | | | | |
RESULT 7
Q94H83 PRELIMINARY; PRT; 748 AA.
ID Q94H83
AC Q94H83;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PUTATIVE HEAT SHOCK PROTEIN.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoideae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Buell C.R., Yuan Q., Ouyang S., Moffat K.S., Hill J.N., Gansberger K.,
RA Brenner M., Burgess S., Hance M., Shvartsbeyn M., Tsitrin T.,
RA Riggs F., Hsiao J., Zismann V., Blunt S., Pai G., Vanaken S.E.,
RA Uterback T.R., Feldblyum T.V., Quackenbush J., Salzberg S.L.,
RA White O., Fraser C.M.;
RT "Oryza sativa chromosome 3 BAC OSUNBA009p23 genomic sequence.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC084380; AAK52144.1; -.
KW Heat shock.
SQ SEQUENCE 748 AA; 83405 MW; 5CF14F897C71B7E2 CRC64;
```

```

Query Match      14.5%; Score 89.5; DB 10; Length 748;
Best Local Similarity 29.5%; Pred. No. 0.8;
Matches 31; Conservative 13; Mismatches 46; Indels 15; Gaps 4;

QY 7 RKALRRARRLAVRMRALESSORPRGSPRSONNIY-----SACPRRARG-ADAAGTGE 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 122 KRALYDOKRKLWLKR--NFSQTNKASAPGASNGFYNAANAASKVTRGNKQKAGPAT 179
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 APVPGCAPLPAPLQVSESPWLHAPSLKTSK-----EYVSLY 98
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 180 SSVRQRPPPPPPRQAPAPPAPKAPPTFTWTSCKNKNYELKVY 224
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 8
Q9RUB1 ID Q9RUB1 PRELIMINARY; PRT; 526 AA.
AC Q9RUB1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE ALGP-RELATED PROTEIN.
GN DRL480.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moifat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Yamathavan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
DR EMBL; AE001992; AAF11045.1; -.
DR TIGR; DRL480; -.
KW Complete proteome.
SQ SEQUENCE 526 AA; 55942 MW; A6C159D4612FE13F CRC64;

Query Match      14.4%; Score 88.5; DB 16; Length 526;
Best Local Similarity 30.2%; Pred. No. 0.69;
Matches 29; Conservative 14; Mismatches 34; Indels 19; Gaps 5;

QY 3 AVERKALRRARRLAVRMRALESSORPRGSPRSONNIYACPRRARGADAAGTGE-- 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 386 AAESLHFSHRAQRI-IGTAWTESIEEPKATPAP-----AAPPAKQTQGTGTATW 436
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 APVPCGPAP-----LPPAPLQVSESPWLHAPSLK 69
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 437 GGSAPAPAPAKGWESTSPAP-KPGDAPQNAPEVK 471
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 9
Q9NKP7 ID Q9NKP7 PRELIMINARY; PRT; 2936 AA.
AC Q9NKP7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE L712.2.
GN L712.2.
OS Leishmania major.
OC Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FRIDLIN;

```

```

RA Myler P.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AC005930; AAF39727.1; -.
SQ SEQUENCE 2936 AA; 305212 MW; B5689E280242FE6B CRC64;

Query Match      14.2%; Score 87.5; DB 5; Length 2936;
Best Local Similarity 35.3%; Pred. No. 5.5;
Matches 30; Conservative 5; Mismatches 37; Indels 13; Gaps 3;

QY 12 RRARLRVEMRALESSORPRGSPRSONNIYAC-----PRRARGADAAGTGEAPV 63
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 77 RKRELTAAATAIARGSPRGSAHAAPAAAPLSALLAPTRHPQRGRTATAAASRTPS 136
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 64 PGGAPLPAPLQVSESPWLHAPSL 88
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 137 PPKSAP-PPTPPPPSSP-----PSL 156
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
Q96UM3 ID Q96UM3 PRELIMINARY; PRT; 821 AA.
AC Q96UM3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE KIAA1802 PROTEIN (FRAGMENT).
GN KIAA1802.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=21245130; PubMed=11347906;
RA Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.;
RA "Prediction of the coding sequences of unidentified human genes. XX.
RA The complete sequences of 100 new cDNA clones from brain which code
RA for large proteins in vitro.";
RL DNA Res. 8:85-95(2001).
DR EMBL; AB058705; BAB47431.1; -.
FT NON_TER 1
SQ SEQUENCE 821 AA; 90139 MW; 7948811E6E9F1DD8 CRC64;

Query Match      14.1%; Score 87; DB 4; Length 821;
Best Local Similarity 31.7%; Pred. No. 1.6;
Matches 33; Conservative 10; Mismatches 45; Indels 16; Gaps 4;

QY 9 ALSRRARLRVEMRALESSORPRGSPRSONNIYACPRRARGADAAGTGEAPVPGP 67
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 304 AVSPEPRRPAVAPGSGWKPGPGSPRPMKSNPSASSGPMKPAKPAPSVS-----PGPW 357
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 58 APLEPAPLQVSESPWLHAPSL-----KTSCEYVSLYHQPAAMME 106
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 358 KPDP-----SVSPGPMKPTPSVSSASMKSSVSPSSWKSPSPAPSE 397
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
Q30801 ID Q30801 PRELIMINARY; PRT; 166 AA.
AC Q30801;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE METHANOL OXIDATION GENES, MXAE, MXAH, MXAB, AND PMI-LIKE GENES
DE (FRAGMENT).
OS Methylobacterium extorquens.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Methylobacterium group; Methylobacterium.
OX NCBI_TaxID=408;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AM1;

```

```

RX MEDLINE=98156126; PubMed=9495022;
RA Springer A.L., Auman A.J., Lidstrom M.E.;
RT "Sequence and characterization of mxaB, a response regulator involved
RT in regulation of methanol oxidation, and of mxaW, a methanol-regulated
RL gene in Methylobacterium extorquens AM1.";
RL FEMS Microbiol. Lett. 160:119-124(1998).
DR EMBL: AF017434; AAC46168.1; -
FT NON_TER
SQ SEQUENCE 166 AA; 17738 MW; E276F6B1E9E1AC67 CRC64;

Query Match 13.9%; Score 85.5; DB 2; Length 166;
Best Local Similarity 32.4%; Pred. No. 0.4;
Matches 33; Conservative 9; Mismatches 35; Indels 25; Gaps 4;

Qy 7 RKALSRARRLAVMRALSS-----SORPRGSPRRS-----QNNIYSACPRRAR 51
Dy 73 RDLLESARRLAASDGGTESTGWNVEVTRTQRDAVTFRSLTSVLPLPGDVQVSPAK-- 130
Qy 52 GADAAGTGEPVPGCAPLPAPLQVSESPWLHAPSLKTSCE 93
Dy 131 ----AGTGEAPAPAPTRGGAAAPGVSA-----APALRSSIE 164

RESULT 12
O77238 PRELIMINARY; PRT; 523 AA.
AC O77238;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE EYEGONE.
GN EYEG OR CG10488.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.

Query Match 13.8%; Score 85; DB 5; Length 523;
Best Local Similarity 25.2%; Pred. No. 1.5;
Matches 30; Conservative 15; Mismatches 42; Indels 32; Gaps 4;

Qy 2 RAVERRKALSRARRLAVMRALSSORPRGSPRRSONNIYSACPRRAGDAAGTGEA 61
Dy 288 RRHQRMNLLKRRSSPNPLHSQSNADAPSPPSNHSASTSAP-----VA 335
Qy 62 PVPGGAPLP-----PAP---LQVSESP-----WLHAPSLKTSCEYVSLYHQ 100
Dy 336 PVPVPPQPLPLCGDHSPOGPPSVLLHPLHGPFGGHHHHHPLHPLTHPAALQLLSHYHQ 394

RESULT 13
Query Match 13.8%; Score 85; DB 5; Length 554;
Best Local Similarity 25.2%; Pred. No. 1.6;
Matches 30; Conservative 15; Mismatches 42; Indels 32; Gaps 4;

Qy 2 RAVERRKALSRARRLAVMRALSSORPRGSPRRSONNIYSACPRRAGDAAGTGEA 61
Dy 288 RRHQRMNLLKRRSSPNPLHSQSNADAPSPPSNHSASTSAP-----VA 335
Qy 62 PVPGGAPLP-----PAP---LQVSESP-----WLHAPSLKTSCEYVSLYHQ 100
Dy 336 PVPVPPQPLPLCGDHSPOGPPSVLLHPLHGPFGGHHHHHPLHPLTHPAALQLLSHYHQ 394

O77238 PRELIMINARY; PRT; 554 AA.
AC O9VTX7;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE EYEG PROTEIN.
GN EYEG OR CG10488.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.-L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balley R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: WITH OTHER HOMEOBOX PROTEINS.
DR EMBL: AE003541; AAF49918.1; -.
DR HSSP; P06601; lFJL.
DR FlyBase; FBgn0000625; eyg.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR001523; Paired_box.
DR Pfam; PF00046; homeobox; 1.
DR Pfam; PF00292; PAX; 1.
DR SMART; SM00389; HOX; 1.
DR SMART; SM00351; PAX; 1.
DR PROSITE; PS00027; HOMEOBOX_1; 1.
DR PROSITE; PS50071; HOMEOBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
SQ SEQUENCE 523 AA; 56495 MW; 9A33F12E86D9F67E CRC64;

Query Match 13.8%; Score 85; DB 5; Length 523;
Best Local Similarity 25.2%; Pred. No. 1.5;
Matches 30; Conservative 15; Mismatches 42; Indels 32; Gaps 4;

Qy 2 RAVERRKALSRARRLAVMRALSSORPRGSPRRSONNIYSACPRRAGDAAGTGEA 61
Dy 288 RRHQRMNLLKRRSSPNPLHSQSNADAPSPPSNHSASTSAP-----VA 335
Qy 62 PVPGGAPLP-----PAP---LQVSESP-----WLHAPSLKTSCEYVSLYHQ 100
Dy 336 PVPVPPQPLPLCGDHSPOGPPSVLLHPLHGPFGGHHHHHPLHPLTHPAALQLLSHYHQ 394

RESULT 13
Query Match 13.8%; Score 85; DB 5; Length 554;
Best Local Similarity 25.2%; Pred. No. 1.6;
Matches 30; Conservative 15; Mismatches 42; Indels 32; Gaps 4;

Qy 2 RAVERRKALSRARRLAVMRALSSORPRGSPRRSONNIYSACPRRAGDAAGTGEA 61
Dy 288 RRHQRMNLLKRRSSPNPLHSQSNADAPSPPSNHSASTSAP-----VA 335
Qy 62 PVPGGAPLP-----PAP---LQVSESP-----WLHAPSLKTSCEYVSLYHQ 100
Dy 336 PVPVPPQPLPLCGDHSPOGPPSVLLHPLHGPFGGHHHHHPLHPLTHPAALQLLSHYHQ 394

O77238 PRELIMINARY; PRT; 523 AA.
AC O77238;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE EYEGONE.
GN EYEG OR CG10488.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.

Query Match 13.8%; Score 85; DB 5; Length 523;
Best Local Similarity 25.2%; Pred. No. 1.5;
Matches 30; Conservative 15; Mismatches 42; Indels 32; Gaps 4;

Qy 2 RAVERRKALSRARRLAVMRALSSORPRGSPRRSONNIYSACPRRAGDAAGTGEA 61
Dy 288 RRHQRMNLLKRRSSPNPLHSQSNADAPSPPSNHSASTSAP-----VA 335
Qy 62 PVPGGAPLP-----PAP---LQVSESP-----WLHAPSLKTSCEYVSLYHQ 100
Dy 336 PVPVPPQPLPLCGDHSPOGPPSVLLHPLHGPFGGHHHHHPLHPLTHPAALQLLSHYHQ 394

O77238 PRELIMINARY; PRT; 554 AA.
AC O9VTX7;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE EYEG PROTEIN.
GN EYEG OR CG10488.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.-L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balley R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: WITH OTHER HOMEOBOX PROTEINS.
DR EMBL; AE003541; AAF49918.1; -.
DR HSSP; P06601; lFJL.
DR FlyBase; FBgn0000625; eyg.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR001523; Paired_box.
DR Pfam; PF00046; homeobox; 1.
DR Pfam; PF00292; PAX; 1.
DR SMART; SM00389; HOX; 1.
DR SMART; SM00351; PAX; 1.
DR PROSITE; PS00027; HOMEOBOX_1; 1.
DR PROSITE; PS50071; HOMEOBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
SQ SEQUENCE 523 AA; 56495 MW; 9A33F12E86D9F67E CRC64;
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: October 28, 2002, 17:28:11 : Search time 16.25 Seconds
 (without alignments)
 586.214 Million cell updates/sec

Title: US-09-135-238B-2
Perfect score: 2055
Sequence: 1 MDRWLWPLYFLPVSGALRIL.....HQPAAMMEDSDSDDYINVPA 390
Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues
Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : Issued Patents.AA.*
 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
 5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*
 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Query %		Length	DB	ID	Description
	Score	Match				
SUMMARIES						
1	2047	99.6	390	3	US-08-961-564A-2	Sequence 2, Appli
2	550	26.8	107	3	US-08-961-564A-4	Sequence 4, Appli
3	198	9.6	771	3	US-08-434-000A-8	Sequence 8, Appli
4	198	9.6	771	4	US-09-312-157-8	Sequence 8, Appli
5	187	9.1	109	3	US-08-961-564A-9	Sequence 9, Appli
6	185	9.0	769	3	US-08-434-000A-10	Sequence 10, Appli
7	185	9.0	769	4	US-09-312-157-10	Sequence 10, Appli
8	182	8.9	608	4	US-09-095-385-4	Sequence 4, Appli
9	182	8.9	746	3	US-08-434-000A-4	Sequence 4, Appli
10	182	8.9	746	4	US-09-312-157-4	Sequence 4, Appli
11	182	8.9	757	3	US-08-434-000A-6	Sequence 6, Appli
12	182	8.9	757	4	US-09-312-157-6	Sequence 6, Appli
13	166.5	8.1	624	2	US-08-642-406A-22	Sequence 22, Appli
14	166.5	8.1	624	4	US-09-199-534-22	Sequence 22, Appli
15	166.5	8.1	773	3	US-08-434-000A-2	Sequence 2, Appli
16	166.5	8.1	773	4	US-09-312-157-2	Sequence 2, Appli
17	117.5	5.7	476	3	US-08-487-550-4	Sequence 8, Appli
18	117	5.7	453	3	US-08-466-151-8	Sequence 8, Appli
19	117	5.7	453	4	US-08-466-163B-8	Sequence 2, Appli
20	113.5	5.5	2337	3	US-08-713-118-2	Sequence 2, Appli
21	113.5	5.5	2337	4	US-09-452-007-2	Sequence 2, Appli
22	112	5.5	451	2	US-08-887-352B-14	Sequence 14, Appli
23	112	5.5	451	2	US-08-887-352B-16	Sequence 16, Appli
24	112	5.5	451	3	US-08-466-151-65	Sequence 65, Appli
25	112	5.5	451	4	US-09-109-207C-14	Sequence 14, Appli
26	112	5.5	451	4	US-09-109-207C-16	Sequence 16, Appli
27	112	5.5	451	4	US-09-296-005-14	Sequence 14, Appli

Sequence 16, Appl
Sequence 10, Appl
Sequence 47, Appl
Sequence 47, Appl
Sequence 6, Appl
Sequence 4, Appl
Sequence 18, Appl
Sequence 16, Appl
Sequence 13, Appl
Sequence 1, Appl
Sequence 18, Appl
Sequence 2, Appl
Sequence 18, Appl
Sequence 1, Appl
Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-08-961-564A-2
; Sequence 2, Application US/08961564A
; Patent No. 6114515
; GENERAL INFORMATION:
; APPLICANT: WU, SHUJIAN
; APPLICANT: SWEET, RAYMOND
; APPLICANT: TRUNEH, ALEMSEGED
; TITLE OF INVENTION: PIGRL-1, A MEMBER OF IMMUNOGLOBULIN
; TITLE OF INVENTION: GENE SUPERFAMILY
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,564A
; FILING DATE: 30-OCT-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/056,935
; FILING DATE: 25-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70236
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 390 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-961-564A-2

Query Match 99.6%; Score 2047; DB 3; Length 390;
Best Local Similarity 99.7%; Pred. No. 2e-178;
Matches 389; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDRWLMPFLPVSGALRLPEVKVEGLGSGVTIKCPLEPMHVRIYLCREMGSGTCTG 60
DB 1 MDRWLMPFLPVSGALRLPEVKVEGLGSGVTIKCPLEPMHVRIYLCREMGSGTCTG 60
QY 61 VVSTNFIKAEYKGRVTLKQYPRKNLFLVEVTQTESDGSVYACGAGMNDRGKTKVTL 120
DB 61 VVSTNFIKAEYKGRVTLKQYPRKNLFLVEVTQTESDGSVYACGAGMNDRGKTKVTL 120
QY 121 NVHSEYPSWEQPMPTPKWFLHPLYLQMPAYASSKFEVTRVTPAQRKVPVPHHSSP 180
DB 121 NVHSEYPSWEQPMPTPKWFLHPLYLQMPAYASSKFEVTRVTPAQRKVPVPHHSSP 180
QY 181 TTQITHRPRVSRASSVAGDKPRTFLPSTTASIKISALEGLLKQPTPSYNNHHTLHQRALD 240
DB 181 TTQITHRPRVSRASSVAGDKPRTFLPSTTASIKISALEGLLKQPTPSYNNHHTLHQRALD 240
QY 241 YGQSQRGQGFHLLIPTILGLFILALLGLVWRAVERKALSRARRLAVRMALESSQ 300
DB 241 YGQSQRGQGFHLLIPTILGLFILALLGLVWRAVERKALSRARRLAVRMALESSQ 300
QY 301 RPRGSRPRSONNTYSACPRRARGADAAGTGEAPVPGAPLPAPLQVSESPWLHAPSL 360
DB 301 RPRGSRPRSONNTYSACPRRARGADAAGTGEAPVPGAPLPAPLQVSESPWLHAPSL 360
QY 361 KTSCEVSYLHOPAAAMDESDSDYINVPA 390
DB 361 KTSCEVSYLHOPAAAMDESDSDYINVPA 390

RESULT 2

US-08-961-564A-4
; Sequence 4, Application US/08961564A
; Patent No. 6114515
; GENERAL INFORMATION:
; APPLICANT: WU, SHUJIAN
; APPLICANT: SWEET, RAYMOND
; APPLICANT: TRUENH, ALENGED
; TITLE OF INVENTION: PIGRL-1, A MEMBER OF IMMUNOGLOBULIN
; TITLE OF INVENTION: GENE SUPERFAMILY
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSO for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,564A
; FILING DATE: 30-OCT-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/056,935
; FILING DATE: 25-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F.
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70236
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 107 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-961-564A-4
Query Match 26.8%; Score 550; DB 3; Length 107;
Best Local Similarity 99.0%; Pred. No. 6.6e-43;
Matches 104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MDRWLMPFLPVSGALRLPEVKVEGLGSGVTIKCPLEPMHVRIYLCREMGSGTCTG 60
DB 1 MDRWLMPFLPVSGALRLPEVKVEGLGSGVTIKCPLEPMHVRIYLCREMGSGTCTG 60
QY 61 VVSTNFIKAEYKGRVTLKQYPRKNLFLVEVTQTESDGSVYACG 105
DB 61 VVSTNFIKAEYKGRVTLKQYPRKNLFLVEVTQTESDGSVYACG 105
RESULT 3
US-08-434-000A-8
; Sequence 8, Application US/08434000A
; Patent No. 6046037
; GENERAL INFORMATION:
; APPLICANT: ANDREW C. HIATT, JULIAN
; APPLICANT: K.-C. MA, THOMAS LEHNER
; TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
; TITLE OF INVENTION: PROTEINS IN PLANTS AND THEIR USES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,000A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below:
; APPLICATION NUMBER: 08/367,395
; FILING DATE: 12/30/94
; ATTORNEY/AGENT INFORMATION:
; NAME: Guise, Jeffrey W.
; REGISTRATION NUMBER: 34,613
; REFERENCE/DOCKET NUMBER: 212/127
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 552-8400
; TELEFAX: (619) 552-0159
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 771 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; TOPOLOGY: DESCRIPTION: Mouse Polyimmunoglobulin Receptor
US-08-434-000A-8
Query Match 9.6%; Score 198; DB 3; Length 771;
Best Local Similarity 31.5%; Pred. No. 1.2e-09;
Matches 53; Conservative 23; Mismatches 68; Indels 24; Gaps 5;
QY 10 FLPSGALRLPEVKVEGLGSGVTIKCPLE-----MHVRIYLCREMGSGTCTGVVST 65

Db 13 FSGVSTKSPFGQVSSIEGDSVSTCYYPDTSVNRHTRKYWCQGA-SGMCITLISN 71
QY 66 NFIKAEYGRVTLKQYPRKNLFLVEVTQLTSDSGVIYACGAGMNTDRGKTQKVTNLNVSE 125
Db 72 GYLSKEYSGRANLINFENNTFVINEQITQDDTGSYKCGLG-TSNRGLSFDVSLEV--- 127
QY 126 YEPSWEEQMPETPKWFHLYPYLFQMPAYASSSKFVTRVTTTPAQRGKVP 173
Db 128 -----SQVPELPSDTHV-----YTKDIGRNVITIECPFKRENVP 160

RESULT 4
US-09-312-157-8
; Sequence 8, Application US/09312157
; Patent No. 6303341
; GENERAL INFORMATION:
; APPLICANT: ANDREW C. HIATT, JULIAN
; K.-C. MA, THOMAS LEHNER
; TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
; PROTEINS IN PLANTS AND THEIR USES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/312,157
; FILING DATE: 14-May-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/434,000
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Guise, Jeffrey W.
; REGISTRATION NUMBER: 34,613
; REFERENCE/DOCKET NUMBER: 212/127
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 552-8400
; TELEFAX: (619) 552-0159
; TELEX: 67-351
; SEQUENCE LISTING
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 771 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; DESCRIPTION: Mouse Polyimmunoglobulin Receptor
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-312-157-8

Query Match 9.6%; Score 198; DB 4; Length 771;
Best Local Similarity 31.5%; Pred. No. 1.2e-09;
Matches 53; Conservative 23; Mismatches 68; Indels 24; Gaps 5;
QY 10 FLPVSGALRLLPEVKEGELGSGVTIKCPLE---MHVRIYLCREMAGSGTCGTGVVSTT 65
Db 13 FSGVSTKSPFGQVSSIEGDSVSTCYYPDTSVNRHTRKYWCQGA-SGMCITLISN 71
QY 66 NFIKAEYGRVTLKQYPRKNLFLVEVTQLTSDSGVIYACGAGMNTDRGKTQKVTNLNVSE 125
Db 72 GYLSKEYSGRANLINFENNTFVINEQITQDDTGSYKCGLG-TSNRGLSFDVSLEV--- 127

QY 126 YEPSWEEQMPETPKWFHLYPYLFQMPAYASSSKFVTRVTTTPAQRGKVP 173
Db 128 -----SQVPELPSDTHV-----YTKDIGRNVITIECPFKRENVP 160
RESULT 5
US-08-961-564A-9
; Sequence 9, Application US/08961564A
; Patent No. 6114515
; GENERAL INFORMATION:
; APPLICANT: WU, SHUIJIAN
; APPLICANT: SWEET, RAYMOND
; APPLICANT: TRUNEH, ALEMSEGED
; TITLE OF INVENTION: PIGRL-1, A MEMBER OF IMMUNOGLOBULIN
; TITLE OF INVENTION: GENE SUPERFAMILY
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,564A
; FILING DATE: 30-OCT-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/056,935
; FILING DATE: 23-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70236
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 109 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-961-564A-9
Query Match 9.1%; Score 187; DB 3; Length 109;
Best Local Similarity 38.3%; Pred. No. 7.7e-10;
Matches 41; Conservative 19; Mismatches 41; Indels 6; Gaps 3;
QY 21 PEVKVEGELGSGVTIKCPLE---EMHVRIYLCREM-AGSGTCGTGVVSTTNFIKAEYKGR 75
Db 4 PE-EVNSVEGNSVITCYYPPTSVNRHTRKYWCQRCARGCLGITLISSEGYVSSKYAGR 62
QY 76 VTLKQYPRKNLFLVEVTQLTSDSGVIYACGAGMNTDRGKTQKVTNLNV 122
Db 63 ANLTNPFENGTFVNVIAQLSQDDSGRYKCGLGINSRLGLSFDVSLEV 109
RESULT 6
US-08-434-000A-10
; Sequence 10, Application US/08434000A
; Patent No. 6046037
; GENERAL INFORMATION:
; APPLICANT: ANDREW C. HIATT, JULIAN
; APPLICANT: K.-C. MA, THOMAS LEHNER

```

1 COUNTRY: U.S.A.
2 ZIP: 90071
3
4 COMPUTER READABLE FORM:
5 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
6 storage
7
8 COMPUTER: IBM Compatible
9 OPERATING SYSTEM: IBM P.C. DOS 5.0
10 SOFTWARE: Word Perfect 5.1
11
12 CURRENT APPLICATION DATA:
13 APPLICATION NUMBER: US/09/312,157
14 FILING DATE: 14-May-1999
15 CLASSIFICATION: <Unknown>
16
17 PRIOR APPLICATION DATA:
18 APPLICATION NUMBER: 08/434,000
19 FILING DATE: <Unknown>
20 ATTORNEY/AGENT INFORMATION:
21 NAME: Guise, Jeffrey W.
22 REGISTRATION NUMBER: 34,613
23 REFERENCE/DOCKET NUMBER: 212/127
24
25 TELECOMMUNICATION INFORMATION:
26 TELEPHONE: (619) 552-8400
27 TELEFAX: (619) 552-0159
28 TELEX: 67-351
29
30 SEQUENCE LISTING
31 INFORMATION FOR SEQ ID NO: 10:
32     SEQUENCE CHARACTERISTICS:
33         LENGTH: 769 amino acids
34         TYPE: amino acid
35         STRANDEDNESS: single
36         TOPOLOGY: linear
37
38     DESCRIPTION: Rat PolyImmunoglobulin Receptor
39
40     SEQUENCE DESCRIPTION: SEQ ID NO: 10:
41
42 US-09-312-157-10
43
44 Query Match          9.0%; Score 185; DB 4; Length 769;
45 Best Local Similarity 34.3%; Pred. No. 1.8e-08;
46 Matches 41; Conservative 21; Mismatches 41; Indels 16; Gaps
47
48 QY 30 GGSVITKCLPE----MHVRIYLCLRMAGSGCGTGVSTTNFIKAYKGRVTLKOYPRKN 85
49      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
50 DB 33 GNSVSTCYTPDTSVNRHTRYKWCROGA-NGYCATLISNGYLSKEYSGRASLINFPPNS 91
51
52 QY 86 LFIWEVTLTSDSGVACAGAMNTDQKTKVTLNVHSEYEPSWEQPMPEPKWPHL 144
53      ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
54 DB 92 TFIINTAHLTQEDTSYKCGLG-TTNRGLFDVSLEV-----SQVPEEPNDTHV 139
55
56 RESULT 8
57 US-09-095-385-4
58     Sequence 4, Application US/09095385
59     Patent No. 6300104
60
61 GENERAL INFORMATION:
62 APPLICANT: Morrison, Sherie L.
63 APPLICANT: Chintalacharuvu, Kote R.
64 TITLE OF INVENTION: SECRETORY IMMUNOGLOBULIN PRODUCED
65 TITLE OF INVENTION: BY SINGLE CELLS AND METHODS FOR MAKING AND USING
66 TITLE OF INVENTION: SAME
67 NUMBER OF SEQUENCES: 4
68 CORRESPONDENCE ADDRESS:
69 ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
70 STREET: 11150 Santa Monica Boulevard, Suite 400
71 CITY: Los Angeles
72 STATE: CA
73 COUNTRY: USA
74 ZIP: 90025
75
76 COMPUTER READABLE FORM:
77 MEDIUM TYPE: diskette
78 COMPUTER: IBM Compatible
79 OPERATING SYSTEM: DOS
80 SOFTWARE: FastSeq for Windows Version 2.0
81
82 CURRENT APPLICATION DATA:
83 APPLICATION NUMBER: US/09/095,385
84 FILING DATE: 09-JUN-1998

```

CLASSIFICATION:
PRIOR APPLICATION DATA: 60/050,969
FILING DATE: 19-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Canady, Karen S
REGISTRATION NUMBER: 39,927
REFERENCE/DOCKET NUMBER: 30435.45USU1
TELEPHONE: 310 445-1140
TELEFAX: 310 445-9031
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 608 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-095-385-4

Query Match 8.9%; Score 182; DB 4; Length 608;
Best Local Similarity 37.4%; Pred. No. 2.4e-08;
Matches 43; Conservative 20; Mismatches 40; Indels 12; Gaps 5;

QY 12 PVSGALRLPEVKVEGELGSGVTIKCPLP-----EMHVRIYLCREMAGSGTCGTVVSTTNF 67
Db 21 PIFG-----PE-EVNSVEGNSVITCYPPPTSVNRHTRKYWCQGGG-CITLISSEGY 73
QY 68 IKAETKGRVTLKQYPRKNLFLVEVTQLTESDSGVYACGAGMNTDRGKTQKVTNLN 122
Db 74 VSSKYAGRANLTNFPENGTFVFNIAQLSQDSDSGRYKCGLGINS-RGLSPFVDSLEV 127

RESULT 9
US-08-434-000A-4
Sequence 4, Application US/08434000A
Patent No. 6046037
GENERAL INFORMATION:
APPLICANT: ANDREW C. HIATT, JULIAN
APPLICANT: K.-C. MA, THOMAS LEHNER
TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
TITLE OF INVENTION: PROTEINS IN PLANTS AND THEIR USES
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/434,000A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below:
APPLICATION NUMBER: 08/367,395
FILING DATE: 12/30/94
ATTORNEY/AGENT INFORMATION:
NAME: Guise, Jeffrey W.
REGISTRATION NUMBER: 34,613
REFERENCE/DOCKET NUMBER: 212/127
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 552-8400

TELEFAX: (619) 552-0159
TELEX: 67-3510
TELEX: SEQUENCE LISTING
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 746 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
TOPOLOGY: DESCRIPTION: Human Polyimmunoglobulin Receptor
US-08-434-000A-4

Query Match 8.9%; Score 182; DB 3; Length 746;
Best Local Similarity 37.4%; Pred. No. 3.2e-08;
Matches 43; Conservative 20; Mismatches 40; Indels 12; Gaps 5;

QY 12 PVSGALRLPEVKVEGELGSGVTIKCPLP-----EMHVRIYLCREMAGSGTCGTVVSTTNF 67
Db 3 PIFG-----PE-EVNSVEGNSVITCYPPPTSVNRHTRKYWCQGGG-CITLISSEGY 55
QY 68 IKAETKGRVTLKQYPRKNLFLVEVTQLTESDSGVYACGAGMNTDRGKTQKVTNLN 122
Db 56 VSSKYAGRANLTNFPENGTFVFNIAQLSQDSDSGRYKCGLGINS-RGLSPFVDSLEV 109

RESULT 10
US-09-312-157-4
Sequence 4, Application US/09312157
Patent No. 6303341
GENERAL INFORMATION:
APPLICANT: ANDREW C. HIATT, JULIAN
APPLICANT: K.-C. MA, THOMAS LEHNER
TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
TITLE OF INVENTION: PROTEINS IN PLANTS AND THEIR USES
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/312,157
FILING DATE: 14-May-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/434,000
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Guise, Jeffrey W.
REGISTRATION NUMBER: 34,613
REFERENCE/DOCKET NUMBER: 212/127
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 552-8400
TELEFAX: (619) 552-0159
TELEX: 67-351
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 746 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: Human Polyimmunoglobulin Receptor
SEQU ID NO: 4

US-09-312-157-4

Query Match 8.9%; Score 182; DB 4; Length 746;
Best Local Similarity 37.4%; Pred. No. 3.2e-08;
Matches 43; Conservative 20; Mismatches 40; Indels 12; Gaps 5;
QY 12 PVSGALRIILPEVKVEGLGGVTIKCPLP-----EMHVRIYLCREMAGSGTGTGVTNNF 67
Db 3 PIFG-----PE-EYNSEGVSSITCYPTTSVNRHTRKYWCROGARGG-CITLISSEGY 55
QY 68 IKAHYGRVTLKQYPRKNFLVEVTQLTESDSGYVACGAGMNTDRGKTOKVTNLN 122
Db 56 VSSKYAGRANLTNFPENGTFVNIQAQLSDSDSGRYKGLGINS-RGLSFDVSLEY 109

RESULT 11

US-08-434-000A-6
; Sequence 6, Application US/08434000A
; Patent No. 6046037
; GENERAL INFORMATION:
; APPLICANT: ANDREW C. HIATT, JULIAN
; APPLICANT: K.-C. MA, THOMAS LEHNER
; TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
; TITLE OF INVENTION: PROTEINS IN PLANTS AND THEIR USES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,000A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION DATA: including application 1
; PRIOR APPLICATION DATA: described below:
; APPLICATION NUMBER: 08/367,395
; FILING DATE: 12/30/94
; ATTORNEY/AGENT INFORMATION:
; NAME: Guise, Jeffrey W.
; REGISTRATION NUMBER: 34,613
; REFERENCE/DOCKET NUMBER: 212/127
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 552-8400
; TELEFAX: (619) 552-0159
; TELEX: 67-3510
; SEQUENCE LISTING
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 757 amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-434-000A-6

Query Match 8.9%; Score 182; DB 3; Length 757;
Best Local Similarity 34.1%; Pred. No. 3.2e-08;
Matches 42; Conservative 21; Mismatches 44; Indels 16; Gaps 4;
QY 17 LRILPEVKVEGL-----GGSVTIKCPDP-----EMHVRIYLCREMAGSGTGTGTV 62
Db 10 LAIFPVVSMKSPIFGPEVSSVEGSRVSIKYYPTTSVNRHTRKYWCROGA-QGRCTTLI 68

QY 63 STTNFIKAHYKGRVTLKQYPRKNFLVEVTQLTESDSGYVACGAGMNTDRGKTOKVTNLN 122
Db 69 SSEGYSDDYVGRANLTNFPESGTFVVDISHLTHKDSGRYKGLGISS-RGLNFDVSLEY 127
QY 123 HSE 125
Db 128 SQD 130

RESULT 12

US-09-312-157-6
; Sequence 6, Application US/09312157
; Patent No. 6303341
; GENERAL INFORMATION:
; APPLICANT: ANDREW C. HIATT, JULIAN
; APPLICANT: K.-C. MA, THOMAS LEHNER
; TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
; TITLE OF INVENTION: PROTEINS IN PLANTS AND THEIR USES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/312,157
; FILING DATE: 14-May-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/434,000
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Guise, Jeffrey W.
; REGISTRATION NUMBER: 34,613
; REFERENCE/DOCKET NUMBER: 212/127
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 552-8400
; TELEFAX: (619) 552-0159
; TELEX: 67-351
; SEQUENCE LISTING
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 757 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; DESCRIPTION: Bovine Polyimmunoglobulin Receptor
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-312-157-6

Query Match 8.9%; Score 182; DB 4; Length 757;
Best Local Similarity 34.1%; Pred. No. 3.2e-08;
Matches 42; Conservative 21; Mismatches 44; Indels 16; Gaps 4;

QY 17 LRILPEVKVEGL-----GGSVTIKCPDP-----EMHVRIYLCREMAGSGTGTGTV 62
Db 10 LAIFPVVSMKSPIFGPEVSSVEGSRVSIKYYPTTSVNRHTRKYWCROGA-QGRCTTLI 68
QY 63 STTNFIKAHYKGRVTLKQYPRKNFLVEVTQLTESDSGYVACGAGMNTDRGKTOKVTNLN 122
Db 69 SSEGYSDDYVGRANLTNFPESGTFVVDISHLTHKDSGRYKGLGISS-RGLNFDVSLEY 127
QY 123 HSE 125


```

; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,000A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below: 1
; APPLICATION NUMBER: 08/367,395
; FILING DATE: 12/30/94
; ATTORNEY/AGENT INFORMATION:
; NAME: Guise, Jeffrey W.
; REGISTRATION NUMBER: 34,613
; REFERENCE/DOCKET NUMBER: 212/127
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 552-8400
; TELEFAX: (619) 552-0159
; TELEX: 67-3510
; TELEX: SEQUENCE LISTING
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 773 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; TOPOLOGY: DESCRIPTION: Rabbit polyimmunoglobulin receptor
; US-08-434-000A-2

Query Match      8.18; Score 166.5; DB 3; Length 773;
Best Local Similarity 42.7%; Pred. No. 8.6e-07;
Matches 44; Conservative 11; Mismatches 41; Indels 7; Gaps 4;

QY 30 GGSVTIKCLPLP---EMHVIYLCREMGSGTGTWVSTTNFIKAYKGRVTLKQYPRKN 85
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 39 GDSVSICTCYPTTSVTRHSRKFWCRE-EESGRCVTLAS-TGYTSQEYSGRGKLTDFPDKG 96
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 86 LFLVETVLTEDSGYACGAGMNTDRGKTQKVTLNHSEYEP 128
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 97 EFVTVTDLTQNDSGSYKGVGVN-GRGLDFGVNVIVSQKPEP 138
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | |

Search completed: October 28, 2002, 17:33:10
Job time : 19.25 secs

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2002, 17:29:31 ; Search time 20.4435 Seconds
(without alignments)
1833.089 Million cell updates/sec

Title: US-09-135-238B-2
Perfect score: 2055
Sequence: 1 MDRWLWPLYFLPVSGALRIL.....HQPAAAMMEDSDDDYINVPA 390

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR.71.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	185	9.0	769	1 QRRTGS	secretory componen
2	182	8.9	757	1 S48841	secretory componen
3	182	8.9	757	2 I45956	polymERIC immunogl
4	182	8.9	764	1 ORHUGS	secretory componen
5	166.5	8.1	773	1 QRRRG	secretory componen
6	117	5.7	1106	1 TVHUGL	transforming prote
7	113.5	5.5	1415	1 EDBEGA	immediate-early pr
8	112.5	5.5	224	2 I37243	CMRF-35 antigen -
9	112	5.5	2082	2 T37056	probable multi-dom
10	111.5	5.4	390	1 QQB877	glycoprotein I pre
11	111.5	5.4	392	2 E88108	protein C46E10.3 l
12	111	5.4	801	2 T29018	hypothetical prote
13	110.5	5.4	2339	2 A42566	omega-conotoxin-se
14	107.5	5.2	528	2 B75310	conserved hypothet
15	107	5.2	1791	2 T02345	hypothetical prote
16	106.5	5.2	460	2 T45968	hypothetical prote
17	106	5.2	496	2 T01564	hypothetical prote
18	103	5.0	478	2 I53960	PRR2 alpha human
19	103	5.0	526	2 D75391	Algp-related prote
20	103	5.0	730	2 S24376	2-aminobenzoyl-CoA
21	102.5	5.0	1008	2 T41244	SEC14 protein homo
22	102	5.0	530	2 A53437	poliovirus recepto
23	101.5	4.9	166	2 PL0012	Ig heavy chain pre
24	101.5	4.9	456	2 T45610	proanthranilate N
25	101.5	4.9	627	2 S14683	Ig mu chain precur
26	101.5	4.9	2152	2 T45583	hypothetical prote
27	101.5	4.9	2237	2 T45115	N-type calcium cha
28	101.5	4.9	3530	2 A59266	unconventional myo
29	101	4.9	475	2 S01321	Ig gamma-2b chain

30	100.5	4.9	504	2 A56205	transcription fact
31	100.5	4.9	876	1 A57988	regulatory protein
32	100.5	4.9	1616	2 I37183	gene APXL protein
33	100	4.9	384	1 I38890	dual specificity p
34	99.5	4.8	2274	2 T30258	adenomatous polypo
35	99	4.8	430	2 I48142	B-lymphocyte antig
36	98.5	4.8	116	2 S04094	transition protein
37	98.5	4.8	847	2 JH0371	B-cell adhesion pr
38	98	4.8	554	1 F0HUMP	macrophage colony-
39	98	4.8	1021	2 T42634	connectin/titin -
40	97.5	4.7	368	2 T51200	hypothetical prote
41	97.5	4.7	870	2 T09137	extensin homolog -
42	97	4.7	1013	2 T46422	hypothetical prote
43	97	4.7	3649	1 S18268	delta-(L-alpha-ami
44	96.5	4.7	1060	2 S63993	acrosomal protein
45	96.5	4.7	2245	2 T27511	hypothetical prote

ALIGNMENTS

RESULT 1

QRRTGS
secretory component precursor - rat
N:Alternate names: poly-Ig receptor; polymeric immunoglobulin receptor
N:Contains: free secretory component; transmembrane secretory component
C:Species: Rattus norvegicus (Norway rat)
C>Date: 07-Sep-1990 #sequence_revision 23-Aug-1996 #text_change 16-Jul-1999
C:Accession: S05407; S54731
R:Banting, G.; Brake, B.; Braghetta, P.; Luzio, J.P.; Stanley, K.K.
FEBS Lett. 254, 177-183, 1989
A:Title: Intracellular targeting signals of polymeric immunoglobulin receptors are h
A:Reference number: S05407; MUID:89378226
A:Accession: S05407
A:Molecule type: mRNA
A:Residues: 1-769 <BAN>
A:Cross-references: EMBL:X15741; NID:g56464; PIDN:CAA33758.1; PID:g56465
C:Complex: monomeric as a transmembrane receptor or free in mucosal secretions; heter
amers: hetero-22-mer composed of one chain of secretory component, one chain of immun
C:Superfamily: secretory component; immunoglobulin homology
C:Keywords: duplication; glycoprotein; immunoglobulin receptor; phosphoprotein; trans
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-769/Product: transmembrane secretory component #status predicted <MAM>
F:19-582/Product: free secretory component #status predicted <MATE>
F:33-112/Domain: immunoglobulin homology <IM1>
F:145-222/Domain: immunoglobulin homology <IM2>
F:250-326/Domain: immunoglobulin homology <IM3>
F:363-442/Domain: immunoglobulin homology <IM4>
F:477-548/Domain: immunoglobulin homology <IM5>
F:644-666/Domain: transmembrane #status predicted <TMM>
F:667-769/Domain: intracellular #status predicted <INT>
F:40-110.56-64.152-220.370-440.384-394.484-546.498-505/Disulfide bonds: #status predi
F:90.135.471/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:488-522/Disulfide bonds: (in Ig-unbound form) #status predicted
F:488/Disulfide bonds: interchain (to IgA alpha-1 chain-192) #status predicted
F:522/Binding site: cysteine (Cys) (covalent) (in Ig-bound form) #status predicted
F:678/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 9.0%; Score 185; DB 1; Length 769;

Best Local Similarity 34.5%; Pred. No. 1.1e-05;

Matches 41; Conservative 21; Mismatches 41; Indels 16; Gaps 4;

QY 30 GGSVTTKCPLEPE----MHVRIYLCRMAGSGTCGTWVTNTFIKARYKGRVTLKQYPRKN 85

Db 33 GNSVSTCYTPDTSVRNRHTRKWCROGA-NGYCATLISSNGYLSKEYSGRASLIINFPENS 91

QY 86 LFLVEVTLQTESDGVYACGAGMNTDRGKTQKVTLVNHSEYBPSEWEEQPMPTPKWFL 144

Db 92 TFVINIAHLTQEDTGSYKCGLG-TTNRGLFFDVSLEV-----SQVPEPNDTHV 139

RESULT 2

S48841


```
QY 222 POTPSYNNHTRL---HQRALDYGSGREGOG----- 251
Db 438 PQ-FSPGAQSCSSDRSPAGSAANTDSGVMTGNAGGSTDLSLDEGPCIAGTGLSLR 496
QY 252 -----PHILIPT-ILGLFL--LALLGLVVKRAVERRKALSR-----ARRLAVR 292
Db 497 RLENRLDLQHLRPTIGRLKPLSLHTGTVSRVPPVSLERSSSSSSSISSAYTVS 556
QY 293 RRALESSORPGSPRPRSONNIYSACPR-----RARGADAAGTGEAPVGP-----GAP 341
Db 557 RSSLASPPPPGSPENGASGLPLMAQOYLLRLARYASARGGTSPTAASLDRIGCLP 616
QY 342 LPP-----APLOVSESPWLHAPSLKTSCEYVSLYHOPAMMEDS 380
Db 617 MPWRRAEYPGYNAGVTRRASDPAQADR-----APARVOREKSLGCVHTPTTVAGG 673
QY 381 DS-DDYI 386
Db 674 QNFDPYL 680

RESULT 7
EDBEGA
immediate-early protein IE175 - Marek's disease virus (strain GA)
N:Alternate names: infected cell protein 4
C:Species: Marek's disease virus
C:Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 16-Jul-1999
C:Accession: A42991
R:Anderson, A.S.; Francesconi, A.; Morgan, R.W.
Virology 189, 657-667, 1992
A:Title: Complete nucleotide sequence of the Marek's disease virus ICP4 gene.
A:Reference number: A42991; MUID:92351564
A:Accession: A42991
A:Molecule type: DNA
A:Residues: 1-1415 <AND>
A:Cross-references: GB:M75729; NID:g330950; PIDN:AAA46111.1; PID:g330951
C:Superfamily: herpesvirus immediate-early protein IE175
C:Keywords: DNA binding; early protein; transcription regulation

Query Match 5.5%; Score 113.5; DB 1; Length 1415;
Best Local Similarity 21.1%; Pred. No. 3.3;
Matches 70; Conservative 37; Mismatches 108; Indels 117; Gaps 14;

QY 104 CGAGMNTDRGKTQKVTNLVHSEYEPS--WEQPMPTPKWFHLPYLFQMPAYASSSKFVT 161
Db 44 CNAAMDDQ-QQMPFPLPVPDWNPNSTDMRPPSPSGPK---KDFCGDLPAPLTSG---P 96
QY 162 RVTPPAQKGVPPVHSSPTTQITHRPR-----VS 191
Db 97 RLTPPSS-GRMSELPHTTSSPRSPRGPTSPSNEHIIISPPRNPFSNTHRNVGHS 155
QY 192 RASSVAGDKPRFLPSTT-----ASKISALEGLLKPTQPSYNNHTRLHRORALDYGSG 246
Db 156 RSPSSSSSSSSSSSSSLVLSSPSSSRSPSPRPRADSSSRPRGRGSRGRSG 215
QY 247 REGOGPHILPTILGLFLLALLGLVVKRAVERRKALSRARRLAVMRALESSORPGSP 306
Db 216 POSKG-----RKA-SPTRKLEDEYLPQRTANRRGG 247
QY 307 RPRSONNIYSACPRRGA-----DAAGT--GEAPVPGCAPLPPA 345
Db 248 RPRGRP-----PKSGRAVORNDIQVTSSSLADTSYDLGSGVWVEVPLPPG----- 295
QY 346 PLOVSESPWL-----HAPSLKTSCEYVSLYHQ 372
Db 296 -----RCWFGGLGHRQALTDSPFIVEAIHR 321

RESULT 8
137243
CMRF-35 antigen - human
C:Species: Homo sapiens (man)
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000
```

```
C:Accession: I37243
R:Jackson, D.G.; Hart, D.N.; Starling, G.; Bell, J.I.
Eur. J. Immunol. 22, 1157-1163, 1992
A:Title: Molecular cloning of a novel member of the immunoglobulin gene superfamily h
A:Reference number: I37243; MUID:92249405
A:Accession: I37243
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-224 <RES>
A:Cross-references: EMBL:X66171; NID:g396169; PIDN:CAA46948.1; PID:g396170
C:Genetics:
A:Gene: CMRF35
```

```
Query Match 5.5%; Score 112.5; DB 2; Length 224;
Best Local Similarity 22.5%; Pred. No. 0.49;
Matches 64; Conservative 33; Mismatches 100; Indels 87; Gaps 12;
```

```
QY 8 LYFLPVSGALRILPEVKVEGELGSGVTIKCPLPEMH--VRIYLCREMGAGSGTCGTVYSTT 65
Db 14 LLLLVPGYFPLSHPMVTAVGPGVSGSLVQCYREKERTLNKWCPRP-PQILRCDKIVETK 72
QY 66 NFIKAERYKRYTLKQYPRKNLFLVEYVQLTSDSGVYACGAGMNTDRGKTOKVTNLVHSE 125
Db 73 G-SAGRRGRVSIIRDSPANUSFTVLTLENLTEDAGTYWCGV----- 112
QY 126 YEPSWEEQPMPTPKW---PHLPYLFQMPAYASSSKFVTRVTPPAQKGVPPVHSSPTT 182
Db 113 -----DTP-WLRDFHDPV-----EVEVSVPFAGTTTA-----SSP-- 142
QY 183 QITHRPRVSRASSVAGDKPTEFLPSTASKISALEGLLKQTPSPYNNHTRLHRORALDYG 242
Db 143 -----QSSMGTSGPTKLPVHTWPSVT-----RKDSEPSPH----- 174
QY 243 SQSGREGQGFILPTILGL-FLIALLGLVVKRAVERRKALSR 285
Db 175 --PGSLFSNVRFLLVLELPLLLMLGAVLWNRQPSRSSRQ 216
```

RESULT 9

```
T37056
probable multi-domain beta keto-acyl synthase - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 28-Jul-2000
C:Accession: T37056
R:Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1999
A:Reference number: 221620
A:Accession: T37056
A:Molecule type: DNA
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Residues: 1-2082 <SEE>
A:Cross-references: EMBL:AL109747; PIDN:CAB52353.1; GSDB:GN00070; SCODEB:SCJ21.07
A:Experimental source: strain A3(2)
C:Genetics:
```

```
A:Gene: SCODEB:SCJ21.07
C:Superfamily: Streptomyces coelicolor probable polyketide synthase; 3-oxoacyl-[acyl-
acyl-carrier-protein] S-malonyltransferase homology
C:Keywords: carrier protein; phosphopantetheine; phosphoprotein
P:48-479/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS>
P:1229/Binding site: phosphopantetheine (Ser) (covalent) #status predicted
```

```
Query Match 5.5%; Score 112; DB 2; Length 2082;
Best Local Similarity 24.1%; Pred. No. 6.6;
Matches 73; Conservative 20; Mismatches 124; Indels 86; Gaps 12;
```

```
QY 105 GAGMNTDRGKTQKVTNLVHSEYEPSWEEQPMPTPKWFHLPYLFQMPAYASSSKFVTRV 163
Db 917 GAGADTFHAALGELAVRGVPLDNLPTPYASPVTPAKEREPRM-----TVK 963
QY 164 TTPAQKGVPP--VHHSSPTTQITHRPRVSRASSVAGDKPTEFLPSTASKISALEGLK 221
Db 964 ISGANYGQLPSPGTPHVEPTPADGPRPDVAPVAAPVAVALPSQAGDV----- 1015
```

```
QY 222 PTPSY-----NHTR-----LHQRALDYGSGREG 249
| | | | | | | | | | | | | | | | | | | | | | | | |
Db 1016 PLPPVYEPAAASSPOEFGSPHAPADARPVIGDDROLAVESIHRTA-ETHLACQRL 1074
| | | | | | | | | | | | | | | | | | | | | | | | |
QY 250 QGFHILITLGLFLALLGLVWRAVERKALSRARRLAVRMRALESSQRGRPR 309
| | | | | | | | | | | | | | | | | | | | | | | | |
Db 1075 TDSHMAFLRMTETTLAALLGAPVPGVPGMTA-----PAALPLPR 1114
| | | | | | | | | | | | | | | | | | | | | | | | |
QY 310 SONNI---YSACPRRARGAAGTGEA-PVFGPGAPLPAPL-----QVSE-SPWLHAPSL 360
| | | | | | | | | | | | | | | | | | | | | | | | |
Db 1115 SVGPLDPAVAPVPAAPAAVGAATPYVPEASAPPASAMRPSVSEPAQAAPAR 1174
| | | | | | | | | | | | | | | | | | | | | | | | |
QY 361 KTS 363
| | | | | | | | | | | | | | | | | | | | | | | | |
Db 1175 DAS 1177
```

RESULT 10

QBE77

glycoprotein I precursor - human herpesvirus 1 (strain 17)
C:Species: human herpesvirus 1
A:Note: host Homo sapiens (man)
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Jul-1999
C:Accession: A05243
R:McGeoch, D.J.; Dolan, A.; Donald, S.; Rixon, F.J.
J. Mol. Biol. 181, 1-13, 1985
A:Title: Sequence determination and genetic content of the short unique region in the genome of human herpesvirus 1
A:Reference number: A06656; MUID:85160822
A:Accession: A05243
A:Molecule type: DNA
A:Residues: 1-390 <MG>
A:Cross-references: GB:L00036; NID:g291490; PIDN:AAA96681.1; PID:g291497
C:Genetics:

Query Match 5.4%; Score 111.5; DB 1; Length 390;
Best Local Similarity 23.2%; Pred. No. 1.1;
Matches 88; Conservative 35; Mismatches 171; Indels 85; Gaps 16;
QY 17 LRILPEYKVEGE-----LGGSVTI-----KCLPEMHVRIYLC--REWAGSGTCGT 60
| | | | | | | | | | | | | | | | | | | | | | | | |
Db 51 LLILGELRFVGDQVPHTTYDGGVELMWHYPMGHKCPRVHVTVTACPRRPAVAFALCRA 110
| | | | | | | | | | | | | | | | | | | | | | | | |
QY 61 VVSTTFIKAEYKGRVTLKQYPRKNLFLVEVTOLTESDGVYA---CGAGMNTD----- 111
| | | | | | | | | | | | | | | | | | | | | | | | |
Db 111 TDSHSS--PAYPLEULMAOP-----LLQVQRATRDYAGVYVLRVWGDAPNASLFVLG 163
| | | | | | | | | | | | | | | | | | | | | | | | |
QY 112 RGKTQKVTLVNHSEYEPESWEQPMPTPKWFLPYLFQMPAYASSKFVTRVTPA-QRG 170
| | | | | | | | | | | | | | | | | | | | | | | | |
Db 164 MAIAEGTLAYNGSAYGSCDKLLPSS-----APRLAPAS-----VYQPAQNA 207
| | | | | | | | | | | | | | | | | | | | | | | | |
QY 171 KVPVHSSPTQITHRPRVSRASSVAGDKPRTFLSTTASKEGGLLKPQTPSYNNH 230
| | | | | | | | | | | | | | | | | | | | | | | | |
Db 208 STPSTTTSTPSTTI-----PAPSTTIAPQASTTTPPTGDPKPPQPGVNH 253
| | | | | | | | | | | | | | | | | | | | | | | | |
QY 231 TRLHQRALDYGSGREGGFGHILIP-TILGLFLALLGLVWRAVERKALSR----- 285
| | | | | | | | | | | | | | | | | | | | | | | | |
Db 254 PPSNATRATRDSRYALVTQIIQTAIPASIIALVFLGSCICFTHRC-QRYRKRSPRIYS 312
| | | | | | | | | | | | | | | | | | | | | | | | |
QY 286 -----ARRLAVRMRALESSQRGRPRPRSONN-----IYSACPRRARGAAGT 330
| | | | | | | | | | | | | | | | | | | | | | | | |
Db 313 PQMPTGISCAVNEAAMARLGAELKSHSTPPKSRSSRTMPSPSLTAIAESEEPAGAAGL 372
| | | | | | | | | | | | | | | | | | | | | | | | |
QY 331 GEAPVPGCAPLPAPLOV 349
| | | | | | | | | | | | | | | | | | | | | | | | |
Db 373 PTPPV-DPTTPTPTPLLV 390
| | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 11

E88108
protein C46E10.3 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: E88108
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating bio
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/projects/C-
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;
A:Accession: E88108
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-392 <STO>
A:Cross-references: GB:chr_II; PIDN:AB96688.1; PID:g2773167; GSPDB:GN00020; CESP:C46E
C:Genetics:
A:Gene: C46E10.3
A:Map position: 2

Query Match 5.4%; Score 111.5; DB 2; Length 392;

Best Local Similarity 23.7%; Pred. No. 1.1;
Matches 81; Conservative 47; Mismatches 121; Indels 93; Gaps 20;

```
QY 88 LVEVTOLTESDGVYACGAGNMTRDGTQKVTLVNHSEYEPESWEQPMPTPKWF-HLPY 146
| | | | | | | | | | | | | | | | | | | | | | | | |
Db 9 LEEVLKLAESIPDVEL-----RGPAGTAVQV--AREPSAE--PLPKLPRVKEEY 56
| | | | | | | | | | | | | | | | | | | | | | | | |
QY 147 LFQMPAYASSKFVTRVTPAQRGKVPVHHSSPTQITHRPRVSRASSVAGDKPRTFLP 206
| | | | | | | | | | | | | | | | | | | | | | | | |
Db 57 FPENSEQASSP-----PVPQILCASPNGPVIH-----GKSLVAGYRPDDRT 99
| | | | | | | | | | | | | | | | | | | | | | | | |
QY 207 STTASKEISALEGLLKPQT-PSYNHTR-LHRO---RALD-YGSOSGREGGFGHILIP--- 257
| | | | | | | | | | | | | | | | | | | | | | | | |
Db 100 ANTTPRLYIPPNYTPANGPLFOAEPRDLREDLLRQIDVFRQOSFQTAQNYQLVLTNR 159
| | | | | | | | | | | | | | | | | | | | | | | | |
QY 258 TILG-----LFLALLGLVWRAVERKALSRARRLA---VRMALESS 299
| | | | | | | | | | | | | | | | | | | | | | | | |
Db 160 TLLSTFDLEQANRLTSQVKSLSSENAVK---EORDRATRAQSLKKNQRLRELFP 216
| | | | | | | | | | | | | | | | | | | | | | | | |
QY 300 ---QRGRPRPRSONNIYSACPRRARGAAGTGEAPVPGCAP---LPP----- 344
| | | | | | | | | | | | | | | | | | | | | | | | |
Db 217 AVEMVNGRPEPPQCSVINFPVSLAISS-----LPSTSIPATILPKIDNLAKVPSF 267
| | | | | | | | | | | | | | | | | | | | | | | | |
QY 345 AFLQYS-----ESPLHAPSLKTSCEVVSILYHQPAAMMEDSDS 382
| | | | | | | | | | | | | | | | | | | | | | | | |
Db 268 APFOISSIPLSEFFLRSPSLQDDLKLAGIKN-----MODKQN 304
| | | | | | | | | | | | | | | | | | | | | | | | |
```

RESULT 12

T29018

hypothetical protein ZK84.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
C:Accession: T29018
R:Kirsten, J.

submitted to the EMBL Data Library, April 1995
A:Description: The sequence of C. elegans cosmid ZK84.

A:Reference number: Z20553
A:Accession: T29018
A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-801 <KIR>

A:Cross-references: EMBL:U23181; PIDN:AAC48204.1; GSPDB:GN00020; CESP:ZK84.1

A:Experimental source: strain Bristol N2; clone ZK84

C:Genetics:

A:Gene: CESP:ZK84.1

A:Map position: 2

A:Introns: 22/2; 45/3; 108/1

C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homolo

Tue Oct 29 09:47:40 2002

```
Matches 56; Conservative 30; Mismatches 100; Indels 52; Gaps 8;
QY 114 KTKVTLNVHSEYEPSEQMPETPKWFLPYLFQMPAYASSSXEVTRVTTTPAORGKYP 173
Db 696 KSRARGSRSSPEPKTKRTPRRRSRSRSPELTRKARLSRRSRAS--SSPETRSTP 753
QY 174 PVHSSPTTQITHRPRVSRAS--SVAGDKPRTFLPSTTA-----SKISALEGLLKPTPSY 227
Db 754 PRHRSVSSPEPAEKSRSSRRRSASPT--KTTSRGRSPSPKPRGLQRSRSR 810
QY 228 NHTRLHRQALDYGSQ-----GREGGFHILPTILGLFLALLGL 270
Db 811 REKTRTTRRRDRSGSQSTSRRRQRSRSRVTRRRRGSGYH----- 853
QY 271 VVKRAVERRKALSRARLAVRMALSSQ--RPRGSPRP--RSONNIYSACPRAR 323
Db 854 --SRSPARQESRTSSRRRGRSTPPTSKKRSRSTSPAPWKRSRASPATHRRSR 909
```

Search completed: October 28, 2002, 17:33:59
Job time : 26.4435 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2002, 17:24:41 ; Search time 11.0081 seconds
(without alignments)
1371.778 Million cell updates/sec

Title: US-09-135-238B-2
Perfect score: 2055
Sequence: 1 MDRWLPLYLPLVSGALRIL.....HQPAAMMEDSDSDYINVPA 390

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	195	9.5	771	1 PIGR_MOUSE	O70570 mus musculu
2	185	9.0	769	1 PIGR_RAT	P15083 rattus norv
3	182	8.9	757	1 PIGR_BOVIN	P81265 bos taurus
4	182	8.9	764	1 PIGR_HUMAN	P01833 homo sapien
5	166.5	8.1	773	1 PIGR_RABIT	P01832 oryctolagus
6	121	5.9	102	1 PIGR_PIG	Q23244 sus scrofa
7	117	5.7	1106	1 GLI1_HUMAN	P08151 homo sapien
8	113.5	5.5	1415	1 ICP4_HSVMG	Q02362 marek's dis
9	112.5	5.5	224	1 CM35_HUMAN	Q08708 homo sapien
10	111.5	5.4	390	1 VGLI_HSV11	P06487 herpes simp
11	109.5	5.3	2339	1 CCAB_HUMAN	Q00975 homo sapien
12	102	5.0	530	1 PVR2_MOUSE	P32507 mus musculu
13	101.5	4.9	3530	1 MY15_HUMAN	Q90kn7 homo sapien
14	101	4.9	2339	1 CCAB_RABIT	Q05152 oryctolagus
15	100.5	4.9	876	1 AREA_EMENI	P17429 emericeella
16	100.5	4.9	1616	1 APXL_HUMAN	Q13796 homo sapien
17	100	4.9	384	1 DUS5_HUMAN	Q16690 homo sapien
18	99	4.8	430	1 CD19_CAVPO	P25917 cavia porce
19	98.5	4.8	116	1 STP2_RAT	P11101 rattus norv
20	98.5	4.8	341	1 FCG2_CAVPO	Q60513 cavia porce
21	98.5	4.8	847	1 CD22_HUMAN	P20273 homo sapien
22	98	4.8	554	1 CSF1_HUMAN	P09603 homo sapien
23	97	4.7	3649	1 ACVS_NOCLA	P27743 nocardia la
24	95.5	4.6	538	1 PVR2_HUMAN	Q92692 homo sapien
25	95.5	4.6	1833	1 ZEP2_HUMAN	P31629 homo sapien
26	95	4.6	862	1 SM4D_HUMAN	Q92854 homo sapien
27	94.5	4.6	275	1 IAP_GVCP	P41436 cydia pomon
28	94.5	4.6	3149	1 TEGU_EBV	P03186 epstein-bar
29	94	4.6	825	1 IL4R_HUMAN	P24394 homo sapien
30	94	4.6	1183	1 DRPL_RAT	P54258 rattus norv
31	94	4.6	1490	1 CRK7_HUMAN	Q90yv4 homo sapien
32	94	4.6	2440	1 NCRL_HUMAN	O75376 homo sapien
33	93.5	4.5	346	1 ST50_YEAST	P25344 saccharomyc

RESULT 1

ID	PIGR_MOUSE	STANDARD:	PRT:	771 AA.
AC	O70570;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	PolymERIC-immunoglobulin receptor precursor (Poly-IG receptor) (PIGR)			
DE	[Contains: Secretory component].			
GN	PIGR.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6 X CBA; TISSUE=Liver;			
EX	MEDLINE=95138517; PubMed=7836758;			
RA	Piskurich J.F., Blanchard M.H., Youngman K.R., France J.A.,			
RA	Kaetzel C.S.;			
RT	"Molecular cloning of the mouse polymeric Ig receptor. Functional			
RT	regions of the molecule are conserved among five mammalian species.";			
RL	J. Immunol. 154:1735-1747(1995).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=L29/SVJ;			
EX	MEDLINE=98072444; PubMed=9409786;			
RA	Martin M.G., Gutierrez E.M., Lam J.T., Li T.W.H., Wang J.;			
RT	"Genomic cloning and structural analysis of the murine polymeric			
RT	receptor (PIGR) gene and promoter region.";			
RL	Gene 201:189-197(1997).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=L29; TISSUE=Liver;			
RA	de Groot N., Vollebregt E., Lee S.H., Verbeet M.P., de Boer H.A.;			
RT	"Molecular cloning and exon-intron organization of the gene encoding			
RT	the murine polymeric immunoglobulin receptor.";			
RL	Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.			
CC	!- FUNCTION: THIS RECEPTOR BINDS POLYMERIC IGA AND IGM AT THE			
CC	BASOLATERAL SURFACE OF EPITHELIAL CELLS. THE COMPLEX IS THEN			
CC	TRANSPORTED ACROSS THE CELL TO BE SECRETED AT THE APICAL SURFACE.			
CC	DURING THIS PROCESS A CLEAVAGE OCCURS THAT SEPARATE THE			
CC	EXTRACELLULAR (KNOWN AS THE SECRETORY COMPONENT) FROM THE			
CC	TRANSMEMBRANE SEGMENT (BY SIMILARITY).			
CC	!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. ALSO SECRETED.			
CC	!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.			
CC	!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE V-TYPE DOMAINS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			

34	93	4.5	355	1	PLK_CHICK	P07354 gallus gall
35	93	4.5	387	1	PIGM_RHOSO	P26698 rhodococcus
36	92.5	4.5	692	1	RFX2_MOUSE	P48379 mus musculu
37	92.5	4.5	1174	1	KPCI_COCHE	O42632 cochliobolu
38	92.5	4.5	1906	1	KMLS_CHICK	P11799 gallus gall
39	92	4.5	283	1	EXTN_SORBI	P24152 sorghum bic
40	92	4.5	475	1	TFEB_MOUSE	Q9R210 mus musculu
41	92	4.5	1638	1	BRM_DROME	P25439 drosophila
42	92	4.5	2150	1	SDC3_CAEEL	P34706 caenorhabdi
43	91.5	4.5	350	1	VGLI_PVRRI	P07646 pseudorhabd
44	91.5	4.5	534	1	CNRD_CAEEL	P41830 caenorhabdi
45	91.5	4.5	872	1	FPL_MYTCO	Q25434 mytilus cor

ALIGNMENTS

```

DR EMBL; U06431; AAA67440.1; -.
DR EMBL; U83434; AAC53585.1; -.
DR EMBL; U83427; AAC53585.1; JOINED.
DR EMBL; U83428; AAC53585.1; JOINED.
DR EMBL; U83429; AAC53585.1; JOINED.
DR EMBL; U83430; AAC53585.1; JOINED.
DR EMBL; U83431; AAC53585.1; JOINED.
DR EMBL; U83432; AAC53585.1; JOINED.
DR EMBL; U83433; AAC53585.1; JOINED.
DR EMBL; U16524; CAA76272.1; -.
DR EMBL; Y16526; CAA76272.1; JOINED.
DR EMBL; Y16527; CAA76272.1; JOINED.
DR EMBL; Y16528; CAA76272.1; JOINED.
DR EMBL; Y16529; CAA76272.1; JOINED.
DR EMBL; Y16530; CAA76272.1; JOINED.
DR EMBL; Y16531; CAA76272.1; JOINED.
DR EMBL; Y16532; CAA76272.1; JOINED.
DR MGD; MGI:103080; PIGR.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF00047; Ig; 5.
DR SMART; SM00409; Ig; 4.
DR SMART; SM00410; Ig_Like; 2.
DR Immunoglobulin domain; Repeat; Transmembrane; Glycoprotein; Signal.
KW SIGNAL 1 18
FT CHAIN 19 771 POLYMERIC-IMMUNOGLOBULIN RECEPTOR.
FT CHAIN 19 611 SECRETORY COMPONENT.
FT DOMAIN 19 645 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 646 668 POTENTIAL.
FT DOMAIN 669 771 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 33 117 IG-LIKE V-TYPE DOMAIN 1.
FT DOMAIN 145 227 IG-LIKE V-TYPE DOMAIN 2.
FT DOMAIN 250 331 IG-LIKE V-TYPE DOMAIN 3.
FT DOMAIN 363 447 IG-LIKE V-TYPE DOMAIN 4.
FT DOMAIN 477 553 IG-LIKE V-TYPE DOMAIN 5.
FT DISULFID 40 110 POTENTIAL.
FT DISULFID 152 220 POTENTIAL.
FT DISULFID 257 324 POTENTIAL.
FT DISULFID 370 440 POTENTIAL.
FT DISULFID 484 546 POTENTIAL.
FT CARBOHYD 90 90 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 147 147 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 170 170 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 206 206 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 420 420 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 471 471 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 159 159 A -> V (IN REF. 1).
FT CONFLICT 396 396 V -> A (IN REF. 1).
FT CONFLICT 620 620 G -> R (IN REF. 1).
SQ SEQUENCE 771 AA; 84998 MW; 78C81302EC710730 CRC64;

Query Match 9.5%; Score 195; DB 1; Length 771;
Best Local Similarity 34.5%; Pred. No. 7e-07;
Matches 48; Conservative 21; Mismatches 54; Indels 16; Gaps 4;

QY 10 FLPSGALRLILEVKEGELGSGVTIKCPLE-----MHVRYLCREMAGSGCGFVSTT 65
Db 13 FSGVSTKSPFGQEVSSIEGDSVSTCYPTDSVNRHTRKYWCQGA-SGMCUTLISN 71
QY 66 NFIAEKYGRVILKQYPRKNLFLVEVTLQSDSGVACAGAGMTDGRKTVTLNVHSE 125
Db 72 GYLSKEYSGRANLINFENNTEVINIEQLTODTDSYKCGLG-TSRGLSFDVSLEV--- 127
QY 126 YEPSWEQMPETPKWFL 144
Db 128 -----SQVPELPDTHV 139

RESULT 2
PIGR_RAT STANDARD; PRT; 769 AA.
ID PIGR_RAT

```

```

AC P15083:
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE PolymERIC-Immunoglobulin receptor precursor (Poly-IG receptor) (PIGR)
GN [Contains: Secretory component].
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=89378226; PubMed=2776882;
RA Banting G., Brake B., Braghetta P., Luzio J.P., Stanley K.K.;
RT "Intracellular targeting signals of polymeric immunoglobulin
RL receptors are highly conserved between species.";
RL FEBS Lett. 254:177-183(1989).
CC -!- FUNCTION: THIS RECEPTOR BINDS POLYMERIC IGA AND IGM AT THE
CC BASOLATERAL SURFACE OF EPITHELIAL CELLS. THE COMPLEX IS THEN
CC TRANSPORTED ACROSS THE CELL TO BE SECRETED AT THE APICAL SURFACE.
CC DURING THIS PROCESS A CLEAVAGE OCCURS THAT SEPARATE THE
CC EXTRACELLULAR (KNOWN AS THE SECRETORY COMPONENT) FROM THE
CC TRANSMEMBRANE SEGMENT.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. ALSO SECRETED.
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE V-TYPE DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X15741; CAA33758.1; -.
DR PIR; S05407; S05407.
DR HSP; P00702; 1JHL.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF00047; Ig; 5.
DR SMART; SM00409; Ig; 4.
DR SMART; SM00410; Ig_Like; 2.
KW Immunoglobulin domain; Repeat; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 18
FT CHAIN 19 769 POLYMERIC-IMMUNOGLOBULIN RECEPTOR.
FT CHAIN 19 611 SECRETORY COMPONENT.
FT DOMAIN 19 643 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 644 666 POTENTIAL.
FT DOMAIN 667 769 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 33 117 IG-LIKE V-TYPE DOMAIN 1.
FT DOMAIN 145 227 IG-LIKE V-TYPE DOMAIN 2.
FT DOMAIN 250 331 IG-LIKE V-TYPE DOMAIN 3.
FT DOMAIN 363 447 IG-LIKE V-TYPE DOMAIN 4.
FT DOMAIN 477 553 IG-LIKE V-TYPE DOMAIN 5.
FT DISULFID 40 110 POTENTIAL.
FT DISULFID 152 220 POTENTIAL.
FT DISULFID 257 324 POTENTIAL.
FT DISULFID 370 440 POTENTIAL.
FT DISULFID 484 546 POTENTIAL.
FT CARBOHYD 90 90 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 206 206 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 471 471 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 769 AA; 84798 MW; 5F849303400255A7 CRC64;

Query Match 9.0%; Score 185; DB 1; Length 769;
Best Local Similarity 34.5%; Pred. No. 3.7e-06;
Matches 41; Conservative 21; Mismatches 41; Indels 16; Gaps 4;

```



```

DR Pfam: PF01688; Herpes_g1; 1.
KW Glycoprotein.
FT CARBOHYD 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 257 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 390 AA; 41369 MW; 39381BD6B5F08C8 CRC64;

Query Match 5.4%; Score 111.5; DB 1; Length 390;
Best Local Similarity 23.2%; Pred. No. 0.38;
Matches 88; Conservative 35; Mismatches 171; Indels 85; Gaps 16;

Qy 17 LRILPEVKEGE-----LGGSVTI-----KCLPLPMHVRIVLC--REMAAGSGTCGT 60
Dy 51 LLILGELFVGQDPHTTYDGGVELMHPMGHKCPVVHVVTTCFRRPVAFAFALCRA 110
Qy 61 VYSTNFKIAEYKGRVTLKQYPRKNLFLVEVTLTESDSGYA-----CGAGMTD----- 111
Dy 111 TDSTHS--PAYTLENLNLAQP-----LLRVQRATRDYAGVYVLRVWVGDPAPNASLFLVLG 163
Qy 112 RGKTQKVTLVNHSYEYFSPWEQMPETPKWFLPFLYLFQMPAYASSKSFVFRVTTPA-ORG 170
Dy 164 MAIAAEGTLAVNSAYSGVSCDPLKLPSS-----APRLAPAS-----VYQAPAPNQA 207
Qy 171 KYPPVHHSSPTQITRPRVRSASSVAGDKPRTFLPSTTASIKSALGGLKLPOTPSYNNH 230
Dy 208 STPSTTSTPSTTI-----PAPSTTIPAPQASTPFPFGDPKPKPPGPGVNH 253
Qy 231 TRLHQRALDYSQSGREGQGFHILIP--TILGLFLLALLGLVKKRAVERKALSR-- 285
Dy 254 PFSNATRATRSYALTVTQIIQIAIPASIIALVFLGSCICFIHRC-QRRYRSRRPIYS 312
Qy 286 -----ARLAVMKRALESSORPGSPRSONN-----IYSACPRRAGADAAGT 330
Dy 313 POMPTGISCANVAARALGAEKLSHSPTPPKSRKRSRTMPSLTAIBESPPAGAAGL 372
Qy 331 GEAPVPGFAPLPAPLQV 349
Dy 373 PTPPV-DPTTPTPTPLLV 390

RESULT 11
CCAB_HUMAN STANDARD; PRT; 2339 AA.
AC Q00975;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Voltage-dependent N-type calcium channel alpha-1B subunit (Calcium
DE channel, L type, alpha-1 polypeptide isoform 5) (Brain calcium channel
DE III) (Biol.).
GN CACNA1B OR CACNL1A5 OR CACH5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS ALPHA-1B-1 AND ALPHA-1B-2).
RC TISSUE=Brain;
RX MEDLINE=9233586; PubMed=1321501;
RA Williams M.E., Brust P.F., Feldman D.H., Patthi S., Smirson S.,
RA Maroufi A., McCue A.F., Velicelebi G., Ellis S.B., Harpold M.M.;
RT "Structure and functional expression of an omega-conotoxin-sensitive
RT human N-type calcium channel.";
RL Science 257:389-395(1992).
RN [2]
RP SEQUENCE OF 1-94 FROM N.A.
RC TISSUE=Lung fibroblast;
RA Kim D.S., Jung H.H., Park S.H., Chin H.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: VOLTAGE-SENSITIVE CALCIUM CHANNELS (VSCC) MEDIATE THE
CC ENTRY OF CALCIUM IONS INTO EXCITABLE CELLS AND ARE ALSO INVOLVED
CC IN A VARIETY OF CALCIUM-DEPENDENT PROCESSES, INCLUDING MUSCLE
CC CONTRACTION, HORMONE OR NEUROTRANSMITTER RELEASE, GENE EXPRESSION,

```

```

CC CELL MOTILITY, CELL DIVISION AND CELL DEATH. THE ISOFORM ALPHA-1B
CC GIVES RISE TO N-TYPE CALCIUM CURRENTS. N-TYPE CALCIUM CHANNELS
CC BELONG TO THE "HIGH-VOLTAGE ACTIVATED" (HVA) GROUP AND ARE BLOCKED
CC BY OMEGA-CONOTOXIN-GVIA (OMEGA-CTX-GVIA) AND BY OMEGA-AGATOXIN-
CC IIA (OMEGA-AGA-IIIA). THEY ARE HOWEVER INSENSITIVE TO
CC DIHYDROPYRIDINES (DHP), AND OMEGA-AGATOXIN-IVA (OMEGA-AGA-IVA).
CC CALCIUM CHANNELS CONTAINING ALPHA-1B SUBUNIT MAY PLAY A ROLE IN
CC DIRECTED MIGRATION OF IMMATURE NEURONS.
CC -!- SUBUNIT: VOLTAGE-DEPENDENT CALCIUM CHANNELS ARE MULTISUBUNIT
CC COMPLEXES, CONSISTING OF ALPHA-1, ALPHA-2, BETA AND DELTA SUBUNITS
CC IN A 1:1:1:1 RATIO. THE CHANNEL ACTIVITY IS DIRECTED BY THE PORE-
CC FORMING AND VOLTAGE-SENSITIVE ALPHA-1 SUBUNIT. IN MANY CASES, THIS
CC SUBUNIT IS SUFFICIENT TO GENERATE VOLTAGE-SENSITIVE CALCIUM
CC CHANNEL ACTIVITY. THE AUXILIARY SUBUNITS BETA AND ALPHA-2/DELTA
CC LINKED BY A DISULFIDE BRIDGE REGULATE THE CHANNEL ACTIVITY.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; ALPHA-1B-1 (SHOWN HERE) AND
CC ALPHA-1B-2; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: ALPHA-1B-1 AND ALPHA-1B-2 ARE EXPRESSED IN THE
CC CENTRAL NERVOUS SYSTEM, BUT NOT IN SKELETAL MUSCLE OR AORTA.
CC -!- DOMAIN: EACH OF THE FOUR INTERNAL REPEATS CONTAINS FIVE
CC HYDROPHOBIC TRANSMEMBRANE SEGMENTS (S1, S2, S3, S5, S6) AND ONE
CC POSITIVELY CHARGED TRANSMEMBRANE SEGMENT (S4). S4 SEGMENTS
CC PROBABLY REPRESENT THE VOLTAGE-SENSOR AND ARE CHARACTERIZED BY A
CC SERIES OF POSITIVELY CHARGED AMINO ACIDS AT EVERY THIRD POSITION.
CC -!- PTM: PHOSPHORYLATED IN VITRO BY CAM-KINASE II, CAPK, PKC AND CGPK
CC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CALCIUM CHANNEL ALPHA-1 SUBUNITS
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M94172; AAA51897.1; -
CC EMBL; M94173; AAA51898.1; -
CC EMBL; U76666; AAC51138.1; -
CC MIM; 601012; -
CC InterPro; IPR002077; Ca_channel.
CC InterPro; IPR002111; Cat_channel_TripL.
CC InterPro; IPR000636; Cat_ion_chan_non_lig.
CC InterPro; IPR001682; Channel_pore_Ca_Na.
CC Pfam; PF00520; ion_trans; 4.
CC PRINTS; PR00167; CACHANNEL.
KW Ionic channel; Transmembrane; Ion transport; Voltage-gated channel;
KW Calcium channel; Glycoprotein; Repeat; Multigene family;
KW Calcium-binding; Phosphorylation; ATP-binding; Alternative splicing.
FT REPEAT 82 359 I.
FT REPEAT 468 712 II.
FT REPEAT 1137 1419 III.
FT REPEAT 1456 1711 IV.
FT DOMAIN 1 95 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 96 114 S1 OF REPEAT I (POTENTIAL).
FT DOMAIN 115 132 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 133 152 S2 OF REPEAT I (POTENTIAL).
FT DOMAIN 153 163 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 164 183 S3 OF REPEAT I (POTENTIAL).
FT DOMAIN 184 187 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 188 206 S4 OF REPEAT I (POTENTIAL).
FT DOMAIN 207 225 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 226 245 S5 OF REPEAT I (POTENTIAL).
FT DOMAIN 246 331 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 332 356 S6 OF REPEAT I (POTENTIAL).
FT DOMAIN 357 482 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 483 501 S1 OF REPEAT II (POTENTIAL).
FT DOMAIN 502 516 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 517 536 S2 OF REPEAT II (POTENTIAL).
FT DOMAIN 537 544 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 545 S3 OF REPEAT II (POTENTIAL).

```

Query Match	Best Local Similarity	Score	DB 1	Length
Matches	83	Conservative	46	Mismatches 135; Indels 151; Gaps 18;
Qy	75	RVTLKQYPRKNLFLVEYVLTQTESDGVYACAGMNTDRGKTQKVTLVNHVSEYFSWEQ	134	
Db	1894	KATLEQTPQAVLRGARVFLRQKSTSLNSGAIONQESGKESV	1943	
Qy	135	MPETPKFHLPL	166	
Db	1944	TQDAPHEARPPLEGRHSTETIPVGRSGALANDVQMSITRRCPDCEPQGLSQGRAASMP	2003	
Qy	167	AQRQKVPVPVHHSSP	222	
Db	2004	RLAAETOPVTDASPMKRSISTLAORPGHGLCSTTPDRP	2048	


```
DR InterPro: IPR001452; SH3.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00612; IQ; 3.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF00784; MYTH4; 2.
DR PRINTS: PR00193; MYOSINHEAVY.
DR ProDom: PD000355; myosin_head; 1.
DR SMART: SM00015; IQ; 2.
DR SMART: SM00242; MYSC; 1.
DR SMART: SM00139; MYTH4; 2.
DR SMART: SM00326; SH3; 1.
DR PROSITE: PS00660; BAND_41_1; FALSE_NEG.
DR PROSITE: PS00661; BAND_41_2; FALSE_NEG.
DR PROSITE: PS00057; BAND_41_3; 1.
DR PROSITE: PS00096; IQ; 3; FALSE_NEG.
DR PROSITE: PS00002; SH3; FALSE_NEG.
KW Myosin: ATP-binding; Actin-binding; Coiled coil; Repeat; SH3 domain;
KW Calmodulin-binding; Disease mutation; Deafness.
FT DOMAIN 1 1887 HEAD OR MOTOR DOMAIN.
FT DOMAIN 1888 2029 NECK OR REGULATORY DOMAIN.
FT DOMAIN 2030 3530 TAIL.
FT DOMAIN 1323 1350 COILED COIL (POTENTIAL).
FT DOMAIN 1792 1799 ACTIN-BINDING (POTENTIAL).
FT DOMAIN 1902 1924 IQ 1.
FT DOMAIN 1925 1954 IQ 2.
FT DOMAIN 1955 1976 IQ 3.
FT DOMAIN 2867 2953 SH3.
FT DOMAIN 3206 3443 BAND 4.1-LIKE.
FT NP_BIND 1315 1322 ATP (POTENTIAL).
FT VARIANT 2111 2113 N -> Y (IN DFN83; FAMILY FROM BENGKALA).
FT VARIANT 2113 2113 /FTID-VAR_010303.
FT VARIANT 2113 2113 I -> F (IN DFN83; INDIAN FAMILY).
FT VARIANT 2113 2113 /FTID-VAR_010304.
SQ SEQUENCE 3530 AA; 395171 MW; 3D103923D4BCBE4A CRC64;

Query Match 4.9%; Score 101.5; DB 1; Length 3530;
Best Local Similarity 21.3%; Pred. No. 29;
Matches 86; Conservative 47; Mismatches 138; Indels 133; Gaps 21;

QY 14 SGALRIIPKVE-----GELGSGVTIKCPLEMHVRIYLCRMAG----- 54
DB 2200 SGAARTLPPTQLEWTATYERKASMALDVGCFNGD-QFSCPVHWSWG-----EEVAGDILRH 2254
QY 55 -----SGCGVTVSTNPIK-AEYKGR-----VTLKQYPRKNLFLVEVTLQTESDS 99
DB 2255 RGLADGWRGTVMKNGVQMAELAGHDYVLDLVDLSDLELRDFPRKSYFIVGTE----- 2308
QY 100 GYVACGAGMNTDRGKTQKVTLNHVSEYEPSEWEE-----QPMPEPKWFHPLPYLFQMPAY 153
DB 2309 -----GPAASRGPRKPVFGN-----SWDSDEDMSTRPQPE-----HMPKVLDSGY 2350
QY 154 ASSSKFVTRVTTPAQRGKVPVHHSSPTQITHRPVRVSRASSVAGDKPRFLPSTTASKI 213
DB 2351 SSHNQDNGTETAGRTA-----TH-----QESDSLGE-----PAVPHKGL 2387
QY 214 SA-LEGILKPTPSYNHHTLHRQALDY-----GSGSGREGQGFH-----ILIPT 258
DB 2388 DCYLDLSLFDLP-VLSYG-DADLEKPTAIYRMKGGGQPGGSGSGTETPRRPEPKPIPG 2445
QY 259 ILGLFLALLGLLVKRAVERKRLSRA-----RRLAVRMRALESSQRPRGSPRPRSONN 313
DB 2446 LDASTLALQQAFTHKQAVLLARGMTLQATALQQOPLSAALRSIPAEPKPAPEAQPTS--- 2502
QY 314 IYSACPRRACDADAGTG--EAPVPGCAPLPAPLQVSESPML 355
DB 2503 -----VGTGPPAKPVLLRATPKPLAPAPLAKAPRL 2532

RESULT 14
ID CCAB_RABIT STANDARD; PRT; 2339 AA.
AC Q05152;
DT 15-JUL-1999 (Rel. 38, Created)
```


CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL: X52491; CAA36731.1; -
DR PIR: S10017; S10017.
DR PDB: 4GAT; 28-JAN-98.
DR PDB: 5GAT; 28-JAN-98.
DR PDB: 6GAT; 28-JAN-98.
DR PDB: 7GAT; 28-JAN-98.
DR TRANSFAC: T02533; -
DR InterPro: IPR000679; ZnF_GATA.
DR Pfam: PF00320; GATA; 1.
DR PRINTS: PR00619; GATAZNFINGER.
DR SMART: SM00401; ZnF_GATA; 1.
DR PROSITE: PS00344; GATA_ZN_FINGER_1; 1.
DR PROSITE: PS01114; GATA_ZN_FINGER_2; 1.
KW Transcription regulation; Activator; DNA-binding; Zinc-finger;
KW Nuclear protein; Nitrate assimilation; 3D-structure.
FT ZN_FING 673 697 GATA-TYPE.
FT DNA_BIND 721 742 H-T-H MOTIF (PROBABLE).
SQ SEQUENCE 876 AA; 94195 MW; 9ADC2273EE536F98 CRC64;

Query Match 4.9%; Score 100.5; DB 1; Length 876;
Best Local Similarity 22.0%; Pred. No. 6.4;
Matches 77; Conservative 47; Mismatches 137; Indels 89; Gaps 17;
QY 96 EDSG--VYACGAGMTDRGKTOKVTLNHSEYEPSEWQPMETPKWFHLPYLFO-MPA 152
Db 515 EDDGQNFSAAGLAMPAGFGDDSIDMNSNMWETSY-----PNSFQSLPA 560
QY 153 YASS-SKFVT-----RVTPA---QRKVPPVHHSSPT-TQTHRPVRSVSSVAGDKPR 202
Db 561 FAAQHRKHVTIGSADMMDTPEWNGGSLGRTHESAASVSEVRNRDQDPRQKIA----R 616
QY 203 TFLPSTASKISALEGLLKQTPSYNHT-----RLHRQRALDYGSSQG 246
Db 617 TSSTPNTAQ-----LLRQSMQNSHTSPNTPESGLNSAAPSASPGGTKNGEQNG 669
QY 247 -----REGQGFHILPTILGLFLALILGLVVKRAVERRRKALSRRARR 288
Db 670 PTTCTNCFTQTTPLWRNPESQP----LCNACGLF-LKLHGVV--RPLSLKTDVIKKNR 722
QY 289 LAVMRALESSQRPGRSPRRSQNNIYSACPRRARGADAAGTGEAPVPGGA----- 340
Db 723 NSANSLAVGSSRSVSKKSARKNSVQVTPPTAPTSSR-AQSNTTSESPAMPSSSGRGSGVV 781
QY 341 PLPAPLQVSESPWLHAPSLKTSCEKVSLEYHQPAAAMMEDSDSDYINVPA 390
Db 782 PIAAAPPK-SSSAATTPGTNCGGAVQVAPKQRORLEKASDVDMAESFS 830

Search completed: October 28, 2002, 17:31:23
Job time : 19.0081 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2002, 17:24:40 ; Search time 24.7419 Seconds
(without alignments)
1059.473 Million cell updates/sec

Title: US-09-135-238B-2_COPY_18_253

Perfect score: 1254

Sequence: 1 RLTPVKVEGLGGSVTKC.....HRQRALDYGSQSGREGQGFH 236

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*

- 1: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1980.DAT.*
- 2: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1981.DAT.*
- 3: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1982.DAT.*
- 4: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1983.DAT.*
- 5: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1984.DAT.*
- 6: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1985.DAT.*
- 7: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1986.DAT.*
- 8: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1987.DAT.*
- 9: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1988.DAT.*
- 10: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1989.DAT.*
- 11: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1990.DAT.*
- 12: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1991.DAT.*
- 13: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1992.DAT.*
- 14: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1993.DAT.*
- 15: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1994.DAT.*
- 16: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1995.DAT.*
- 17: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1996.DAT.*
- 18: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1997.DAT.*
- 19: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1998.DAT.*
- 20: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1999.DAT.*
- 21: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA2000.DAT.*
- 22: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1254	100.0	390	20	AA1980.DAT.*
2	1254	100.0	390	20	AA1981.DAT.*
3	1254	100.0	390	20	AA1982.DAT.*
4	659	52.6	422	22	AAE05349
5	459	36.6	107	20	AA1980.DAT.*
6	319	25.4	97	21	AA1981.DAT.*
7	197	15.7	771	17	AA1982.DAT.*
8	194	15.5	771	22	AA1983.DAT.*
9	185	14.8	769	17	AA1984.DAT.*
10	185	14.8	769	22	AA1985.DAT.*
11	180.5	14.4	607	20	AA1986.DAT.*

12	180.5	14.4	607	20	AA1987.DAT.*
13	180.5	14.4	746	17	AA1988.DAT.*
14	180.5	14.4	764	22	AA1989.DAT.*
15	180.5	14.4	764	22	AA1990.DAT.*
16	179.5	14.3	757	17	AA1991.DAT.*
17	179	14.3	757	22	AA1992.DAT.*
18	176	14.0	733	22	AA1993.DAT.*
19	175.5	14.0	532	19	AA1994.DAT.*
20	166.5	13.3	584	12	AA1995.DAT.*
21	166.5	13.3	773	17	AA1996.DAT.*
22	166.5	13.3	773	22	AA1997.DAT.*
23	156	12.4	584	12	AA1998.DAT.*
24	126	10.0	211	21	AA1999.DAT.*
25	121	9.6	313	22	AA2000.DAT.*
26	119.5	9.5	203	22	AA2001.DAT.*
27	119.5	9.5	205	20	AA2002.DAT.*
28	117.5	9.4	334	19	AA2003.DAT.*
29	111.5	8.9	307	19	AA2004.DAT.*
30	108.5	8.7	268	22	AA2005.DAT.*
31	106	8.5	298	20	AA2006.DAT.*
32	106	8.5	301	20	AA2007.DAT.*
33	105	8.4	164	20	AA2008.DAT.*
34	103.5	8.3	332	20	AA2009.DAT.*
35	103.5	8.3	332	21	AA2010.DAT.*
36	103.5	8.3	332	21	AA2011.DAT.*
37	103.5	8.3	332	21	AA2012.DAT.*
38	103.5	8.3	332	22	AA2013.DAT.*
39	103.5	8.3	332	22	AA2014.DAT.*
40	100.5	8.0	504	22	AA2015.DAT.*
41	100.5	8.0	504	22	AA2016.DAT.*
42	98.5	7.9	230	22	AA2017.DAT.*
43	98.5	7.9	231	19	AA2018.DAT.*
44	98.5	7.9	260	21	AA2019.DAT.*
45	98	7.8	493	22	AA2020.DAT.*

ALIGNMENTS

RESULT 1	
AA1980.DAT.*	
ID	AA1980.DAT.* standard; Protein: 390 AA.
AC	AA1980.DAT.*
AC	AA1980.DAT.*
DT	17-DEC-1999 (first entry)
DE	Human Toso protein sequence.
DE	Human Toso protein sequence.
KW	Human: Toso protein; target; drug screening; diagnosis; apoptosis;
KW	apoptosis related disease.
OS	Homo sapiens.
PN	WO9950671-A2.
PN	WO9950671-A2.
PD	07-OCT-1999.
PF	30-MAR-1999; 99WO-US06945.
PR	30-MAR-1998; 98US-0050861.
PR	(RIGE-) RIGEL PHARM INC.
PI	Payan D;
PI	WPI; 1999-591379/50.
DR	N-PSDB; AA225422.
XX	Screening agents useful for modulating apoptosis and controlling
PT	apoptosis related diseases
XX	Claim 1; Fig 2a; 75pp; English.
PS	

XX The present invention describes a method of Screening for a bioactive
CC agent capable of binding a Toso protein. Also described a methods for:
CC (1) screening a bioactive agent capable of modulating activity of a
CC Toso cell-surface receptor, comprising adding a candidate bioactive
CC agent to a cell comprising a recombinant Toso nucleic acid, and
CC exposing the cells to an apoptotic agent that will induce apoptosis;
CC (2) modulating apoptosis comprising administering an exogenous
CC compound that binds Toso, to a cell; (3) identifying a cell containing
CC a mutant Toso gene, comprising determining it's sequence; (4)
CC identifying the Toso genotype, comprising determining the sequence of
CC at least one Toso gene; and (5) diagnosing an apoptosis related
CC condition, comprising measuring activity of Toso in a tissue, and
CC comparing to the activity from non-affected individual's tissue, where
CC a reduced activity of the patient indicates risk of an apoptosis related
CC condition. The methods are useful for identifying agents capable of
CC diagnosing and treating apoptosis related disease, their use for
CC modulating apoptosis, and methods for diagnosing the disease state.
CC The present sequence represents the human Toso protein for use in
CC methods from the invention.

XX Sequence 390 AA;

Query Match 100.0%; Score 1254; DB 20; Length 390;
Best Local Similarity 100.0%; Pred. No. 9.3e-111;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILPEVKVEGELGSGVTIKCPLPEMHVRIYLCREMAGSGTCGTVSTTFNFIKAEYKGRVT 60
|||||
Db 18 RILPEVKVEGELGSGVTIKCPLPEMHVRIYLCREMAGSGTCGTVSTTFNFIKAEYKGRVT 77
|||||

QY 61 LKQYPRKNLFLVEVTQLTESDSGVYACGAGMNTDRGKTOKVTLNVHSEYEPSEWEEQPMPE 120
|||||
Db 78 LKQYPRKNLFLVEVTQLTESDSGVYACGAGMNTDRGKTOKVTLNVHSEYEPSEWEEQPMPE 137
|||||

QY 121 TPWFHLPYLFQMPAYASSSKFVTRVTPAQRGKVPVPHHSSPTQTTHRPVRSASSVA 180
|||||
Db 138 TPWFHLPYLFQMPAYASSSKFVTRVTPAQRGKVPVPHHSSPTQTTHRPVRSASSVA 197
|||||

QY 181 GDKPRTELPSTTASKISALEGLLKPOTPSYNHHTRLHQRALDYGSGSGREGQGFH 236
|||||
Db 198 GDKPRTELPSTTASKISALEGLLKPOTPSYNHHTRLHQRALDYGSGSGREGQGFH 253
|||||

RESULT 2
AAI17496
ID AAI17496 standard; Protein; 390 AA.
AC AAI17496;
XX
XX 03-AUG-1999 (first entry)
DT
DE Human Toso protein.
XX
XX Toso protein; tumour necrosis factor mediated apoptosis inhibition;
KW TNF mediated apoptosis; T cell overactivity; autoimmune disease;
KW Sjogrens connective tissue disorder; transplant rejection; cancer.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH Peptide 1..17
FT /label= signal
FT Protein 18..390
FT /label= Toso
FT Region 254..272
FT /label= transmembrane_region
XX
XX W09525832-A1.
XX
XX 27-MAY-1999.
PD
XX 16-NOV-1998; 98WO-US24391.
PF

XX 17-AUG-1998; 98US-0135238.
PR 17-NOV-1997; 97US-0066063.
XX
PA (STRD) UNIV LELAND STANFORD JUNIOR.
XX
PI Hitoshi Y, Nolan GP;
XX
XX WPI; 1999-338007/28.
DR N-PSDB; AAX76123.
XX
PT DNA encoding Toso, a protein having inhibitory effects on TNF
mediated apoptosis
XX
PS Claim 20; Fig 2a; 70pp; English.
XX
XX The present sequence is a Toso protein (I). (I) has anti-apoptotic
CC and cytostatic activity. Toso (named after a Japanese liquor that is
CC drunk on New Year's Day to celebrate long life and eternal youth) most
CC likely acts by induction of cFIP expression which inhibits caspase-8
CC processing. Recombinant (I) can be used to modulate apoptosis in a cell
CC or to treat an apoptosis related condition in a mammal. Apoptosis
CC related conditions can also be treated by administration of the Toso
CC protein or antibody. Apoptosis related or mediated conditions that can
CC be treated include diseases characterized by T cell overactivity, e.g.
CC Sjogrens connective tissue disorder, autoimmune diseases, diseases where
CC T cells actively destroy cells, including transplant rejection and
CC conditions where cells of any kind that are not dying express Toso
CC appropriately, e.g. cancer of T or B cell origin (where increased
CC apoptosis would be appropriate).

XX Sequence 390 AA;

Query Match 100.0%; Score 1254; DB 20; Length 390;
Best Local Similarity 100.0%; Pred. No. 9.3e-111;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RILPEVKVEGELGSGVTIKCPLPEMHVRIYLCREMAGSGTCGTVSTTFNFIKAEYKGRVT 60
|||||
Db 18 RILPEVKVEGELGSGVTIKCPLPEMHVRIYLCREMAGSGTCGTVSTTFNFIKAEYKGRVT 77
|||||

QY 61 LKQYPRKNLFLVEVTQLTESDSGVYACGAGMNTDRGKTOKVTLNVHSEYEPSEWEEQPMPE 120
|||||
Db 78 LKQYPRKNLFLVEVTQLTESDSGVYACGAGMNTDRGKTOKVTLNVHSEYEPSEWEEQPMPE 137
|||||

QY 121 TPWFHLPYLFQMPAYASSSKFVTRVTPAQRGKVPVPHHSSPTQTTHRPVRSASSVA 180
|||||
Db 138 TPWFHLPYLFQMPAYASSSKFVTRVTPAQRGKVPVPHHSSPTQTTHRPVRSASSVA 197
|||||

QY 181 GDKPRTELPSTTASKISALEGLLKPOTPSYNHHTRLHQRALDYGSGSGREGQGFH 236
|||||
Db 198 GDKPRTELPSTTASKISALEGLLKPOTPSYNHHTRLHQRALDYGSGSGREGQGFH 253
|||||

RESULT 3
AAI05001
ID AAI05001 standard; Protein; 390 AA.
XX
XX AAI05001;
AC
XX 16-JUN-1999 (first entry)
DT
DE Human PIGRL-1 protein sequence.
XX
XX PIGRL-1; human; autoimmune disease; hyper-IgM Immunodeficiency; HIM;
KW x-linked Severe Combined Immunodeficiency; XSCID; IgA deficiency;
KW diagnosis; therapy.
XX
XX Homo sapiens.
OS
XX EP905238-A2.
PN
XX 31-MAR-1999.
PD

XX PF 14-AUG-1998; 98EP-0306487.
XX PR 30-OCT-1997; 97US-0961564.
XX PR 25-AUG-1997; 97US-0056935.
XX PA (SMIK) SMITHKLINE BEECHAM CORP.
XX PI Sweet RW, Truneh A, Wu S;
XX PI WPI; 1999-192666/17.
XX DR N-PSDB; AAX28178.
XX PT New polypeptides encoding PIGRL-1 useful for treating diseases such
XX PT as X-linked Severe Combined Immunodeficiency
XX PS Claim 11; Page 7; 26pp; English.
XX CC This sequence is the human PIGRL-1 protein of the invention.
XX CC Autoimmune diseases involving altered expression or activity of PIGRL-1
XX CC may include Hyper-IgM Immunodeficiency (HIM), X-linked Severe Combined
XX CC Immunodeficiency (XSCID) and IgA deficiency. These diseases can be
XX CC diagnosed or susceptibility to them predicted by: (1) determining whether
XX CC there is a mutation in the genomic copy of the gene encoding PIGRL-1; or
XX CC (2) measuring the amount of PIGRL-1 in a sample derived from the patient.
XX CC Patients deficient in PIGRL-1 can be treated by administering either the
XX CC PIGRL-1 DNA or its complement or an agonist of PIGRL-1 to the patient.
XX CC Patients with excessive expression or activity of PIGRL-1 can be treated
XX CC by administering an antagonist of PIGRL-1, an antisense nucleic acid
XX CC molecule which inhibits the expression of PIGRL-1 or administering
XX CC sufficient PIGRL-1 to compete with the endogenous activity. PIGRL-1 can
XX CC be used to identify its agonists by contacting a cell expressing PIGRL-1
XX CC with a candidate compound in the presence of a signal system and noting
XX CC the candidate as an agonist if a signal is produced. The same method can
XX CC be used to identify antagonists of PIGRL-1 but the presence of an
XX CC antagonist is indicated by a decrease in production of the signal.
XX CC Antibodies against PIGRL-1 may be used to isolate or identify clones
XX CC expressing PIGRL-1. Polynucleotides encoding PIGRL-1 may be used to
XX CC identify chromosomal mutations in the gene encoding PIGRL-1 in patients.
XX CC This information may then be correlated with the incidence of autoimmune
XX CC disease in those patients to identify whether the mutation causes the
XX CC disease.
XX SQ Sequence 390 AA;
Query Match 100.0%; Score 1254; DB 20; Length 390;
Best Local Similarity 100.0%; Pred. No. 9.3e-111;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RILPEVKVEGELGGSVTIKCPLEPMHVRIYLCREMAGSGTCGTVVSTTNFKAEYKGRVT 60
DB 18 RILPEVKVEGELGGSVTIKCPLEPMHVRIYLCREMAGSGTCGTVVSTTNFKAEYKGRVT 77
QY 61 LKQYPRKNLFLVEVTQLTESDSGVYACAGMNTDRGKTQKVTNLNVHSEYEPSWEQPMPE 120
DB 78 LKQYPRKNLFLVEVTQLTESDSGVYACAGMNTDRGKTQKVTNLNVHSEYEPSWEQPMPE 137
QY 121 TPKNFHLPLFQMPAYASSKSFVTRVTPAQRGVKVPVHHSSPTQTTHRPRVSRASSVA 180
DB 138 TPKNFHLPLFQMPAYASSKSFVTRVTPAQRGVKVPVHHSSPTQTTHRPRVSRASSVA 197
QY 181 GDKPRTEFLPSTASKISALEGLLPQTPSYNHHTRLHRQALDYGSGSGREGQGQFH 236
DB 198 GDKPRTEFLPSTASKISALEGLLPQTPSYNHHTRLHRQALDYGSGSGREGQGQFH 253
RESULT 4
ID AAE05349
XX AAE05349 standard; Protein; 422 AA.
XX AC AAE05349;
XX DT 12-SEP-2001 (first entry)

XX DE Mouse Toso protein.
XX KW Mouse; cytostatic; antiinflammatory; immunoregulatory; tissue integrity;
KW wound healing; immune response; vaccine; cancer; asthma; allergy;
KW cell trafficking; therapy; secreted protein; Fas-induced apoptosis;
KW Toso.
XX OS Mus sp.
XX PN WO200148192-A1.
XX PD 05-JUL-2001.
XX PF 21-DEC-2000; 2000WO-NZ00256.
XX PR 23-DEC-1999; 99US-0171678.
XX PR 28-NOV-2000; 2000US-0724864.
XX (GENE-) GENESIS RES & DEV CORP LTD.
XX Watson JD, Murison JG;
XX WPI; 2001-425665/45.
XX DR N-PSDB; AAD10117.
XX PT Novel isolated polypeptide useful to isolate corresponding interacting
XX PT proteins or other compounds, to quantitatively determine levels of
XX PT interacting proteins or other compounds, and as therapeutic target -
XX PS Claim 6; Page 78-79; 101pp; English.
XX CC The patent discloses novel polynucleotides and their corresponding
XX CC proteins which play a major role in induction of growth, cell migration
XX CC and proliferation, cell-cell interaction and the differentiation of
XX CC tissue-specific cells. These proteins are important in the maintenance
XX CC of tissue integrity and thus are important in wound healing. They are
XX CC useful in various assays to determine the biological activity, to raise
XX CC antibodies, to isolate corresponding interacting proteins or other
XX CC compounds, to quantitatively determine levels of interacting proteins or
XX CC other compounds, and as therapeutic target in a whole range of disease
XX CC states. Compositions comprising the novel proteins of the invention are
XX CC useful for treating mammalian disorders. Polynucleotides of the invention
XX CC are useful in genome and physical mapping, in positional cloning of
XX CC genes, to tag or identify an organism or its reproductive material (as
XX CC non-disruptive tags for marking organisms), and for the diagnosis and
XX CC treatment of mammalian diseases which is the consequence of inappropriate
XX CC expression of kinase genes. They are useful for promoting immune response
XX CC as part of a vaccine or anti-cancer treatment, as target for cancer
XX CC treatment, as immunoregulatory and anti-inflammatory molecule, as
XX CC diagnostic for specific types of cancer and for development of an
XX CC anti-cancer treatment, and as a target for antagonists in the treatment
XX CC of diseases such as asthma and allergy. They are also useful to inhibit
XX CC or enhance the activity of the soluble molecule that binds proteins of
XX CC the invention, for tissue and neural regeneration, to promote or block
XX CC cell trafficking, and as anti-inflammatory and/or vaccine adjuvant.
XX CC The present sequence is Toso, a secreted protein from mouse. Toso is
XX CC a cell surface, specific regulator of Fas-induced apoptosis in T-cells.
XX SQ Sequence 422 AA;
Query Match 52.6%; Score 659; DB 22; Length 422;
Best Local Similarity 53.7%; Pred. No. 3.5e-54;
Matches 130; Conservative 34; Mismatches 70; Indels 8; Gaps 4;
QY 1 RILPEVKVEGELGGSVTIKCPLEPMHVRIYLCREMAGSGTCGTVVSTTNFKAEYKGRVT 60
DB 18 RVLPEVQLNVWEGSIIIECPQLHVRMYLCRQMAKPGICSTVVSNT-FVKKEYERRY 76
QY 61 LKQYPRKNLFLVEVTQLTESDSGVYACAGMNTDRGKTQKVTNLNVHSEY-EPSWEQPMPE 119
DB 77 LTPCLDKLFLVEMTQLTENDDGIYACGVGMKDKGKTQKITLNVHNEYPEPFWEDEWTS 136

CC disease.
XX Sequence 107 AA:
SQ Query Match 36.6%; Score 459; DB 20; Length 107;
Best Local Similarity 100.0%; Pred. No. 5.7e-36;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RILPEVKVGEGLGSGVTIKCPPEMHVRIYLCREMGSGTGTWSTTFNFIKAEYKGRVT 60
Db 18 RILPEVKVGEGLGSGVTIKCPPEMHVRIYLCREMGSGTGTWSTTFNFIKAEYKGRVT 77
QY 61 LKQYPRKNLFLVEVTQLTESDSGYIACG 88
Db 78 LKQYPRKNLFLVEVTQLTESDSGYIACG 105
RESULT 6
AAV65401
ID AAY65401 standard; Protein; 97 AA.
XX
AC AAY65401;
XX
DT 01-FEB-2000 (first entry)
DE Human 5' EST related polypeptide SEQ ID NO:1562.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; diagnosis;
KW gene therapy; chromosome mapping; upstream regulatory sequence;
KW forensic; location; development; protein synthesis; stability;
KW regulation; identification.
XX
OS Homo sapiens.
XX
PN WO953051-A2.
XX
PD 21-OCT-1999.
XX
PF 09-APR-1999; 99WO-1B00712.
XX
PR 09-APR-1998; 98US-0057719.
PR 28-APR-1998; 98US-0069047.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI: 2000-038446/03.
DR N-PSDB; AAZ43015.
XX
PT Novel secreted protein 5' expressed sequence tag sequences used in
PT diagnostic, forensic, gene therapy, and chromosome mapping procedures
XX
PS Claim 3; Page 820-821; 837pp; English.
XX
CC AAZ42265 to AAZ43075 represent novel 5' expressed sequence tag (EST)
CC sequences, corresponding to human secreted proteins. AAY64651 to
CC AAY65438 represent the EST-related proteins corresponding to AAZ42265 to
CC AAZ43052. The 5' ESTs can be used for producing secreted human gene
CC products. They can be used to identify and isolate 5' untranslated
CC regions (UTRs) and upstream regulatory regions which control the
CC location, development stage, rate, and quantity of protein synthesis, as
CC well as stability of mRNA. The ESTs are also useful as probes for
CC chromosome mapping, and to obtain full length cDNA clones. The ESTs can
CC also be used in forensic procedures to identify individuals, or in
CC diagnostic procedures to identify individuals having genetic diseases
CC resulting from abnormal gene expression. The products may also be used in
CC gene therapy protocols. The nucleic acids encoding signal peptides can be
CC used for directing extracellular secretion of a polypeptide or the
CC insertion of a polypeptide into a membrane, or importing a polypeptide
CC into a cell. The proteins encoded by the EST sequences may be useful in
CC treating a variety of human conditions. Secreted proteins have
CC therapeutic value, and the identification of new secreted proteins is

QY 120 ETPKWFHLPYLFPMP-----AYASSKFTVTVTPAQRGKVPVPHHSSPTQITHRPRVS 174
Db 137 ERPRWLHRFLQHQPWLHSGSEHPSGGVIAKVTTTFAKSTKAPPVHQFSSITSVTQHPRVY 196
QY 175 RASSVAGDKPRTFLPSTTAKSISALEGLKAPQTPPSNHHTRLHRORALDYGSQSGREGQG 234
Db 197 RAFVSATKSPALLPATTASKTSTQQA-IRPLEASYSHHTRLHQRTRRHGHGPHYGRDRG 255
QY 235 FH 236
Db 256 LH 257
RESULT 5
AAV05002
ID AAY05002 standard; peptide; 107 AA.
XX
AC AAY05002;
XX
DT 16-JUN-1999 (first entry)
DE Human PIGRL-1 protein sequence fragment.
XX
KW FIGRL-1; human; autoimmune disease; hyper-IgM Immunodeficiency; HIM;
KW X-linked Severe Combined Immunodeficiency; XSCID; IgA deficiency;
KW diagnosis; therapy.
XX
OS Homo sapiens.
XX
PN EP905238-A2.
XX
PD 31-MAR-1999.
XX
PF 14-AUG-1998; 98EP-0306487.
XX
PR 30-OCT-1997; 97US-0961564.
PR 23-AUG-1997; 97US-0056935.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI Sweet RW, Truneh A, Wu S;
XX
DR WPI: 1999-192666/17.
DR N-PSDB; AAX28179.
XX
PT New polypeptides encoding PIGRL-1 useful for treating diseases such
PT as X-linked Severe Combined Immunodeficiency
XX
PS Disclosure; Page 9; 26pp; English.
XX
CC This sequence is a fragment of the human PIGRL-1 of the invention.
CC Autoimmune diseases involving altered expression or activity of PIGRL-1
CC may include Hyper-IgM Immunodeficiency (HIM), X-linked Severe Combined
CC Immunodeficiency (XSCID) and IgA deficiency. These diseases can be
CC diagnosed or susceptibility to them predicted by: (1) determining whether
CC there is a mutation in the genomic copy of the gene encoding PIGRL-1; or
CC (2) measuring the amount of PIGRL-1 in a sample derived from the patient.
CC Patients deficient in PIGRL-1 can be treated by administering either the
CC PIGRL-1 DNA or its complement or an agonist of PIGRL-1 to the patient.
CC Patients with excessive expression or activity of PIGRL-1 can be treated
CC by administering an antagonist of PIGRL-1, an antisense nucleic acid
CC molecule which inhibits the expression of PIGRL-1 or administering
CC sufficient PIGRL-1 to compete with the endogenous activity. PIGRL-1 can
CC be used to identify its agonists by contacting a cell expressing PIGRL-1
CC with a candidate compound in the presence of a signal system and noting
CC the candidate as an agonist if a signal is produced. The same method can
CC be used to identify antagonists of PIGRL-1 but the presence of an
CC antagonist is indicated by a decrease in production of the signal.
CC Antibodies against PIGRL-1 may be used to isolate or identify clones
CC expressing PIGRL-1. Polynucleotides encoding PIGRL-1 may be used to
CC identify chromosomal mutations in the gene encoding PIGRL-1 in patients.
CC This information may then be correlated with the incidence of autoimmune
CC disease in those patients to identify whether the mutation causes the

New ligands binding to a specific region of a polymeric immunoglobulin receptor, useful for transporting therapeutic or diagnostic compositions into or across cells expressing pIgR e.g. in drug delivery

Disclosure; Fig 1; 102pp; English.

The invention provides ligands that bind specifically to a region of an animal cell polymeric immunoglobulin receptor (pIgR). The pIgR cleaves to produce a stalk region remaining attached to the cell and a secretory component existing in the organ of interest in several forms. The ligands do not bind to the stalk or the most abundant form of the secretory component present in the organ under physiological conditions. The ligands are useful for transporting therapeutic or diagnostic compositions into or across cells expressing pIgR, useful to introduce or transport ligands such as antibodies and/or to deliver biologically active components such as proteins, nucleic acids or detectable labels. They are used to deliver therapeutic compositions to mucosal surfaces. Such as the gastro-intestinal tract, respiratory system etc. in humans. They are also useful to label cells expressing pIgR, e.g. to distinguish epithelial cells from a mixed cell population in pathology studies or to aid in carcinoma diagnosis (since pIgR expression is reduced in carcinomas relative to normal epithelium). They can also be used to deliver veterinary compositions, especially in mammals such as farm, domestic or wild mammals or birds e.g. birds reared for human consumption. The present sequence represents a mouse pIgR sequence.

Query Match 15.5%; Score 194; DB 22; Length 771;

Best Local Similarity 36.1%; Pred. No. 1.2e-09;

Matches 43; Conservative 20; Mismatches 40; Indels 16; Gaps 4;

QY 13 GGSVTIKCPLE-----MHVRIYLCREMAGSGTCGTVSTNFIKAEYKGRVTKQYPRKN 68

DB 33 GDSVSIITCYPTDSVNRHTRKYWCROGA-SGMCTLLISSNGYLSKSGRANLINFENN 91

QY 69 LFLVEVTLQTESDGVYACGAGMNTDRGKTQKVTLNHSEYEPSEWQPMPTPKWFHL 127

DB 92 TFVINIQLTQDTSYKCGLG-TSNRGLSFDVSLEV-----SQVPELPDTHV 139

RESULT 9

AAW03181

ID AAW03181 standard; Protein; 769 AA.

AC AAW03181;

DT 24-FEB-1997 (first entry)

XX Rat poly-immunoglobulin receptor.

DE Rat; immunoglobulin receptor; protection protein; mutants;
 KW heavy chain; antigen binding domain; protection; pathogen;
 KW mucosal; environment; gastrointestinal; passive; immunisation;
 KW Guy's 13 antibody; prevention; dental caries; Streptococcus;
 KW poly; sorbinus.

OS Rattus rattus.

XX Key Location/Qualifiers
 FH Region 13..45

FT /note= "putative immunoglobulin binding residues
 of domain I"

FT Domain 1..120

FT /label= domain_I

FT Domain 110..230

FT /label= domain_II

FT Domain 210..340

FT /label= domain_III

FT Domain 320..450

FT /label= domain_IV

FT Domain 440..550

FT /label= domain_v
 FT 550..606
 FT /note= "external portions of domain VI"
 FT 550..627
 FT /note= "external portions of domain VI"
 FT 625..660
 FT /label= transmembrane_segment
 FT 650..769
 FT /label= intracellular_portion

XX WO9621012-A1.

XX 11-JUL-1996.

XX 27-DEC-1995; 95WO-US16889.

XX 04-MAY-1995; 95US-0434000.

PR 30-DEC-1994; 94US-0367395.

XX (PLAN-) PLANT BIOTECHNOLOGY INC.

PA (UNNE-) UNITED MEDICAL & DENTAL SCHOOLS GUYS.

PA (PLAN-) PLANET BIOTECHNOLOGY INC.

XX Hiatt AC, Lehner T, Ma JKC;

PI WPI; 1996-333987/33.

DR N-PSDB; AAT31291.

XX Immunoglobulin and protection protein complex and its prodn. in plants - useful for passive immunisation against mucosal antigens, esp. against S. mutans and S. sorbinus to prevent dental caries

PS Disclosure; Pages 123-127; 152pp; English.

XX The present sequence is the rat poly-immunoglobulin (Ilg) receptor, a portion of which corresp. to residues 1-627, pref. 1-606 or esp. residues 13-45, 1-120, 110-230, 210-340, 320-450, 440-550, 550-606 or 550-627 comprises a protection protein (PP). The Ig of the invention comprises a PP as above in association with an Ig derived heavy chain, having at least a portion of an antigen (Ag) binding domain. The PP protects the Ig in harsh mucosal, e.g. gastrointestinal, environments, therefore enhancing its effectiveness in passively immunising animals against mucosal pathogens. The Ag binding domain is specifically derived from the Guy's 13 antibody, and the Ig can be used to prevent dental caries by binding, e.g. Streptococcus mutans serotypes c, e and f, or S. sorbinus serotypes d and g.

XX Sequence 769 AA;

Query Match 14.8%; Score 185; DB 17; Length 769;

Best Local Similarity 34.5%; Pred. No. 8.5e-09;

Matches 41; Conservative 21; Mismatches 41; Indels 16; Gaps 4;

QY 13 GGSVTIKCPLE-----MHVRIYLCREMAGSGTCGTVSTNFIKAEYKGRVTKQYPRKN 68

DB 33 GNSVSIITCYPTDSVNRHTRKYWCROGA-NGYCATLISSNGYLSKSGRANLINFENN 91

QY 69 LFLVEVTLQTESDGVYACGAGMNTDRGKTQKVTLNHSEYEPSEWQPMPTPKWFHL 127

DB 92 TFVINIQLTQDTSYKCGLG-TTNRGLSFDVSLEV-----SQVPELPDTHV 139

RESULT 10

AAG65697

ID AAG65697 standard; protein; 769 AA.

XX AAG65697;

XX 07-JAN-2002 (first entry)

XX Rat polymeric immunoglobulin receptor (pIgR) sequence.

XX

KW PolymERIC immunoglobulin receptor; pIgR; ligand; therapeutic;
XX carcinoma diagnosis; veterinary; rat.
OS Rattus sp.

XX WO200172846-A2.

XX PD 04-OCT-2001.

XX PF 26-MAR-2001; 2001WO-US09699.

XX PR 27-MAR-2000; 2000US-192197P.

XX PR 27-MAR-2000; 2000US-192198P.

XX PA (REGC) UNIV CALIFORNIA.

XX PI Mostov KE, Chapin SJ, Richman-Eisenstat J;

XX WI; 2001-611619/70.

XX New ligands binding to a specific region of a polymeric immunoglobulin
XX receptor, useful for transporting therapeutic or diagnostic
XX compositions into or across cells expressing pIgR e.g. in drug delivery
XX
XX Disclosure; Fig 1; 102pp; English.

XX The invention provides ligands that bind specifically to a region of an
XX animal cell polymeric immunoglobulin receptor (pIgR). The pIgR cleaves
XX to produce a stalk region remaining attached to the cell and a secretory
XX component existing in the organ of interest in several forms. The ligands
XX do not bind to the stalk or the most abundant form of the secretory
XX component present in the organ under physiological conditions. The
XX ligands are useful for transporting therapeutic or diagnostic
XX compositions into or across cells expressing pIgR, useful to introduce
XX or transport ligands such as antibodies and/or to deliver biologically
XX active components such as proteins, nucleic acids or detectable labels.
XX They are used to deliver therapeutic compositions to mucosal surfaces
XX such as the gastro-intestinal tract, respiratory system etc. in humans.
XX They are also useful to label cells expressing pIgR e.g. to distinguish
XX epithelial cells from a mixed cell population in pathology studies or to
XX aid in carcinoma diagnosis (since pIgR expression is reduced in
XX carcinomas relative to normal epithelium). They can also be used to
XX deliver veterinary compositions, especially in mammals such as farm,
XX domestic or wild mammals or birds e.g. birds reared for human
XX consumption. The present sequence represents a rat pIgR sequence.

XX Sequence 769 AA;

Query Match 14.8%; Score 185; DB 22; Length 769;
Best Local Similarity 34.5%; Pred. No. 8.5e-09;
Matches 41; Conservative 21; Mismatches 41; Indels 16; Gaps 4;

QY 13 GGSVTIKCLPPE----MHVRIYLCREMAGSGTCGTVTNFIKAEYKGRVTLKQYPRKN 68

Db 33 GNSVITCYYPDTSVNRHTRKYWCROGA-NGYCATLISSNGLYSKEYSGRASLINFPENS 91

QY 69 LFLVETVLTQTESDGVYACGAGMTDRCKTKVTLNVHSEYEPSPWEEQMPETPKWPHL 127

Db 92 TFVINIAHLTQEDTGSYKCGLG-TTNRGLFFDVSLEV-----SQVPEFPNDTHV 139

RESULT 11

ID AAY34099 standard; Protein; 607 AA.

XX AAY34099;

XX 20-DEC-1999 (first entry)

DE Partial amino acid sequence of plasmid pSHUSC.

XX Multimeric protein; immunoglobulin; receptor-ligand complex;

KW hetero-dimeric receptor; trimeric G protein; transgenic.

XX Synthetic.

XX PN WO9949024-A2.

XX PD 30-SEP-1999.

XX PF 24-MAR-1999; 99WO-US06506.

XX PR 25-MAR-1998; 98US-0079249.

XX PA (PLAN-) PLANET BIOTECHNOLOGY INC.

XX PI Wycoff KL, Jaiswal SK;

XX WI; 1999-580446/49.

XX DR N-PSDB; AA222290.

XX Producing heterologous multimeric proteins in plants, transformed with
XX several plasmids expressing polypeptide components, particularly for
XX immunoglobulins -
XX Example 1; Fig 8; 42pp; English.

XX The invention relates to a method for producing heterologous, multimeric
XX proteins in plant cells. The method comprises: (a) transforming the cells
XX with several naked plasmids each encoding some, but not all, of the
XX polypeptide components of the multimeric proteins, and together providing
XX all the polypeptide components; and (b) culturing the cells. The method
XX is used to produce biologically active multimeric proteins particularly
XX immunoglobulins, receptor-ligand complexes, homo- or hetero-dimeric
XX receptors, or trimeric G proteins. This method provides properly
XX associated and assembled multimeric proteins in a fast and efficient
XX process, without the need to cross plants expressing single component of
XX the protein. Transgenic plants containing adjacent and stably integrated
XX plasmids, and their progeny can also express the multimeric proteins. The
XX present sequence represents the partial amino acid sequence of the
XX plasmid pSHUSC.

XX Sequence 607 AA;

Query Match 14.4%; Score 180.5; DB 20; Length 607;
Best Local Similarity 38.7%; Pred. No. 1.7e-08;
Matches 41; Conservative 19; Mismatches 39; Indels 7; Gaps 4;

QY 4 PEVKVEGELGGSVTIKCLP-----EMHVRIYLCREMAGSGTCGTVTNFIKAEYKGRV 59

Db 25 PE-EVNSVEGNSVITCYYPDTSVNRHTRKYWCROGARGG-CITLISSEGYVSSKYAGRA 82

QY 60 TLKQYPRKNLFLVEVTQTESDGVYACGAGMTDRCKTKOKVTLNV 105

Db 83 NLINFPEGTFTVFNIAQLSODSDSGRYKCGLGINS-RGLSFDVSLEV 127

RESULT 12

AAW95601

ID AAW95601 standard; Protein; 607 AA.

XX AAW95601;

XX 08-JUN-1999 (first entry)

DE Human secretory Immunoglobulin A component.

KW Immunoglobulin A; secretory; component; IgA; human; treatment;

KW prevention; infection; HIV; AIDS; cold; flu; virus;

KW human immunodeficiency virus; respiratory syncytial virus.

XX Homo sapiens.

XX WO9857993-A1.

XX

PD 23-DEC-1998.

XX 10-JUN-1998; 98WO-US11975.

XX 19-JUN-1997; 97US-0050969.

XX (REGC) UNIV CALIFORNIA.

XX Chintalacharuvu KR, Morrison SL;

PI WPI; 1999-080950/07.

DR N-PSDB; AAX07407.

XX Producing secretory immunoglobulin in single cells - useful to

PT produce commercial quantities of secretory immunoglobulin to prevent

PT or treat infections

XX Disclosure; Pages 22-24; 39pp; English.

XX The sequence is that of the secretory component of human secretory

CC immunoglobulin A (sigA). It can be used as part of a method for

CC the production of sig molecules. This method is useful for

CC producing commercial quantities of sig (especially sigA) to treat

CC or prevent infections. In particular, sigA produced by the method

CC can be used to prevent or treat infections in mammals, birds or

CC fish; especially systemic infections or infections at a mucosal

CC surface. It is especially useful to prevent or treat infection

CC with human immunodeficiency virus (HIV), respiratory syncytial

CC virus, flu virus or cold virus. The method allows production of

CC commercial quantities of sig molecules for therapeutic use, not

CC previously possible; production using non-plant cells and a

CC single cell type is more efficient than a previous multi-step

CC process of fusing recombinant plant cells, and avoids alterations

CC of the sig by plant cells. SigA molecules are more stable

CC and resistant to proteolysis than previously used IgA molecules,

CC and can be administered to prevent as well as to treat infections,

CC unlike e.g. IgG and IgM molecules.

XX Sequence 607 AA;

SQ

Query Match 14.4%; Score 180.5; DB 20; Length 607;

Best Local Similarity 38.7%; Pred. No. 1.7e-08;

Matches 41; Conservative 19; Mismatches 39; Indels 7; Gaps 4;

QY 4 PEVKVEGELGSGVTIKCPPL----EMHRYIYLCREMGSGTGVVSTTFNFIKAEYKGRV 59

DB 25 PE-EVNSVEGNSVITCYPPTSVNRHTRKYWCRCRGARG-CITLISSEGVSSKYAGRA 82

QY 60 TLKQYPRKNLFLVEVTLTSDSGSVYACGAGMNTDRGKTQKVTLNV 105

DB 83 NLTFNPENGTFVNVIAQLSQDSDSGRYKGLGINS-RGLSFDVSLV 127

RESULT 13

AAW03178

ID AAW03178 standard; Protein; 746 AA.

XX

AC AAW03178;

XX

XX 24-FEB-1997 (first entry)

XX Human poly-immunoglobulin receptor.

XX Human; immunoglobulin; receptor; protection protein; mutants;

KW heavy chain; antigen binding domain; protection; pathogen;

KW mucosal; environment; gastrointestinal; passive; immunisation;

KW Guy's 13 antibody; prevention; dental caries; Streptococcus;

XX poly; sorbinus.

OS Homo sapiens.

XX

XX Key Location/Qualifiers

FT Region 13..45

FT

FT

FT Domain

FT 1..120

FT /label= domain_I

FT 110..230

FT /label= domain_II

FT 210..340

FT /label= domain_III

FT 320..450

FT /label= domain_IV

FT 440..550

FT /label= domain_V

FT 550..606

FT /note= "external portions of domain VI"

FT 550..627

FT /note= "external portions of domain VI"

FT 625..660

FT /label= transmembrane_segment

FT 650..746

FT /label= intracellular_portion

XX WO9621012-A1.

XX 11-JUL-1996.

XX 27-DEC-1995; 95WO-US16889.

XX 04-MAY-1995; 95US-0434000.

XX 30-DEC-1994; 94US-0367395.

XX (PLAN-) PLANT BIOTECHNOLOGY INC.

PA (UNME-) UNITED MEDICAL & DENTAL SCHOOLS GUYS.

PA (PLAN-) PLANET BIOTECHNOLOGY INC.

XX Hiatt AC, Lehner T, Ma JKC;

XX WPI; 1996-333987/33.

DR N-PSDB; AAT31288.

XX Immunoglobulin and protection protein complex and its prodn. in

PT plants - useful for passive immunisation against mucosal antigens,

PT esp. against S. mutans and S. sorbinus to prevent dental caries

XX

XX Disclosure; Pages 105-108; 152pp; English.

XX The present sequence is the human poly-immunoglobulin (Ig)

CC receptor, a portion of which corresp. to residues 1-627, pref.

CC 1-606, or esp. residues 13-45, 1-120, 110-230, 210-340, 320-450,

CC 440-550, 550-606 or 550-627 comprises a protection protein (PP).

CC The Ig of the invention comprises a PP as above in association with

CC an Ig derived heavy chain, having at least a portion of an antigen

CC (Ag) binding domain. The PP protects the Ig in harsh mucosal, e.g.

CC gastrointestinal, environments, therefore enhancing its

CC effectiveness in passively immunising animals against mucosal

CC pathogens. The Ag binding domain is specifically derived from the

CC Guy's 13 antibody, and the Ig can be used to prevent dental caries

CC by binding, e.g. Streptococcus mutans serotypes c, e and f, or

CC S. sorbinus serotypes d and g.

XX

SQ Sequence 746 AA;

Query Match 14.4%; Score 180.5; DB 17; Length 746;

Best Local Similarity 38.7%; Pred. No. 2.2e-08;

Matches 41; Conservative 19; Mismatches 39; Indels 7; Gaps 4;

QY 4 PEVKVEGELGSGVTIKCPPL----EMHRYIYLCREMGSGTGVVSTTFNFIKAEYKGRV 59

DB 7 PE-EVNSVEGNSVITCYPPTSVNRHTRKYWCRCRGARG-CITLISSEGVSSKYAGRA 64

QY 60 TLKQYPRKNLFLVEVTLTSDSGSVYACGAGMNTDRGKTQKVTLNV 105

DB 65 NLTFNPENGTFVNVIAQLSQDSDSGRYKGLGINS-RGLSFDVSLV 109

PR 27-MAR-2000; 2000US-192197P.
PR 27-MAR-2000; 2000US-192198P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
XX Mostov KE, Chapin SJ, Richman-Eisenstat J;
XX WPI; 2001-611619/70.
XX
XX New ligands binding to a specific region of a polymeric immunoglobulin
PT receptor, useful for transporting therapeutic or diagnostic
PT compositions into or across cells expressing pigR e.g. in drug delivery
PT
XX
XX Disclosure; Fig 2; 102pp; English.
XX
XX The invention provides ligands that bind specifically to a region of an
CC animal cell polymeric immunoglobulin receptor (pigR). The pigR cleaves
CC to produce a stalk region remaining attached to the cell and a secretory
CC component existing in the organ of interest in several forms. The ligands
CC do not bind to the stalk or the most abundant form of the secretory
CC component present in the organ under physiological conditions. The
CC ligands are useful for transporting therapeutic or diagnostic
CC compositions into or across cells expressing pigR, useful to introduce
CC or transport ligands such as antibodies and/or to deliver biologically
CC active components such as proteins, nucleic acids or detectable labels.
CC They are used to deliver therapeutic compositions to mucosal surfaces.
CC such as the gastro-intestinal tract, respiratory system etc. in humans.
CC They are also useful to label cells expressing pigR, e.g. to distinguish
CC epithelial cells from a mixed cell population in pathology studies or to
CC aid in carcinoma diagnosis (since pigR expression is reduced in
CC carcinomas relative to normal epithelium). They can also be used to
CC deliver veterinary compositions, especially in mammals such as farm,
CC domestic or wild mammals or birds e.g. birds reared for human
CC consumption. The present sequence represents a human pigR sequence.
XX
SQ Sequence 764 AA:

Query Match 14.4%; Score 180.5; DB 22; Length 764;
Best Local Similarity 38.7%; Pred. No. 2.3e-08;
Matches 41; Conservative 19; Mismatches 39; Indels 7; Gaps 4;

Qy 4 PEVKYEGELGGSVTIKCPLP-----EMHVRIYLORENAGSGTCGTWSTTFIKAEYKGRV 59
Db 25 PE-EVNSVEGNSVITCYYPPTSVMNRHTRKIWCQARGG-CITLISSEGYSSKYAGRA 82

Qy 60 TLKQYPRKNLFLVEVTLQTESDSGVYACGAGMNTDRGKTQKVTNLV 105
Db 83 NLTFNFPENGTFVNIQLSQDDSGRYKGLGINS-RGLSFDVSLEV 127

Search completed: October 28, 2002, 17:30:54
Job time : 27.7419 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2002, 17:24:41 ; Search time 35.121 Seconds
(without alignments)
1921.018 Million cell updates/sec

Title: US-09-135-238B-2

Perfect score: 2055

Sequence: 1 MDRWLWPLFLPVSGALRIL.....HQRAMMEDSDSDYINVPA 390

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2047	99.6	390	060667	Q0667 homo sapien
2	1157	56.3	422	11 Q9D8T1	Q9d8t1 mus musculus
3	186	9.1	534	4 Q96SA2	Q96sa2 homo sapien
4	184	9.0	758	6 Q9N2H7	Q9n2h7 sus scrofa
5	175	8.5	455	11 Q920L8	Q920l8 mus musculus
6	175	8.5	535	11 Q9EQT7	Q9eqt7 mus musculus
7	161	7.8	307	11 Q54947	Q54947 rattus norv
8	143.5	7.0	299	4 Q9UBK4	Q9ubk4 homo sapien
9	143.5	7.0	299	4 Q9UGN4	Q9ugn4 homo sapien
10	126	6.1	359	4 Q4356	Q4356 homo sapien
11	124	6.0	298	4 Q9HD97	Q9hd97 homo sapien
12	122.5	6.0	301	4 Q95100	Q95100 homo sapien
13	118	5.7	820	4 Q60585	Q60585 homo sapien
14	117	5.7	897	11 Q70495	Q70495 mus musculus
15	116.5	5.7	364	4 Q96D42	Q96d42 homo sapien
16	114.5	5.6	335	13 Q9YGV5	Q9ygv5 gallus gall

17	113.5	5.5	1417	12	067631	067631 gallid herp
18	113	5.5	1537	5	Q9VAI2	Q9vai2 drosophila
19	112	5.5	2082	2	Q9S200	Q9s200 streptomyce
20	111.5	5.4	392	5	Q44716	Q44716 caenorhabdi
21	111	5.4	367	10	Q9AYC9	Q9ayc9 oryza sativ
22	111	5.4	801	5	Q23635	Q23635 caenorhabdi
23	109.5	5.3	335	13	Q9PWR4	Q9pwr4 gallus gall
24	109.5	5.3	862	11	Q9JIK1	Q9jik1 rattus norv
25	109.5	5.3	2321	12	Q9DGT6	Q9dt6 turkey herp
26	109	5.3	666	11	Q9NML2	Q9nml2 mus musculu
27	108.5	5.3	923	4	Q9NVB8	Q9nvb8 homo sapien
28	108	5.3	892	4	Q9Y438	Q9y438 homo sapien
29	107.5	5.2	528	16	Q9RSJ1	Q9rsj1 deinococcus
30	107	5.2	270	4	Q9UMT1	Q9umt1 homo sapien
31	107	5.2	270	4	Q9H564	Q9h564 homo sapien
32	107	5.2	355	2	Q93RI1	Q93rl1 streptococc
33	107	5.2	385	2	Q54913	Q54913 streptococc
34	107	5.2	1275	4	Q9UQ36	Q9uq36 homo sapien
35	107	5.2	1783	4	O15038	O15038 homo sapien
36	107	5.2	1791	4	O60382	O60382 homo sapien
37	107	5.2	1847	5	Q9KN5	Q9kn5 leishmania
38	107	5.2	2296	4	Q9UHA8	Q9uha8 homo sapien
39	107	5.2	2752	4	Q9UQ35	Q9uq35 homo sapien
40	106.5	5.2	335	13	Q9YGH1	Q9ygh1 gallus gall
41	106.5	5.2	460	10	Q9LFA8	Q9lfa8 arabidopsis
42	106	5.2	390	4	Q96T50	Q96t50 homo sapien
43	106	5.2	390	4	Q96AP7	Q96ap7 homo sapien
44	106	5.2	496	10	O23094	O23094 arabidopsis
45	106	5.2	800	4	Q96PZ4	Q96pz4 homo sapien

ALIGNMENTS

RESULT 1

060667 PRELIMINARY; PRT: 390 AA.
AC O60667;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ANTI-FAS-INDUCED APOPTOSIS (REGULATOR OF FAS-INDUCED APOPTOSIS).
GN TOSO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98246048; PubMed=9586636;
RA Hitoshi Y., Lorens J., Kitada S.I., Fisher J., Labarge M., Ring H.Z.,
RA Francke U., Reed J.C., Kinoshita S., Nolan G.P.;
RT "Toso, a cell surface, specific regulator of Fas-induced apoptosis in
T cells.";
RL Immunity 8:461-471(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=PRIMARY B-CELLS FROM TONSILS;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL: AF057557; AAC18830.1; -
DR EMBL: BC006401; AAH06401.1; -
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00409; Ig; 1.
SQ SEQUENCE 390 AA; 43146 MW; FE91D217EBCA99C6 CRC64;

Query Match 99.6%; Score 2047; DB 4; Length 390;
Best Local Similarity 99.7%; Pred. No. 2e-174;
Matches 389; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MDRWLWPLFLPVSGALRILPEVKVEGELGSGSVTIKCPLPEMHVRIYLCREMAGSGTCCT 60

Db 1 MDRWLPYFLPVSGALRILPEVKVEGELGSGVTIKCPLEMHVRIYLCREMGSGTCCT 60
Qy 61 VSTTNFKAIEYKGRVTLKQYPRKNIPLVEVTOLTESDSGVYACGAGMNTDRGKTOKVTL 120
Db 61 VSTTNFKAIEYKGRVTLKQYPRKNIPLVEVTOLTESDSGVYACGAGMNTDRGKTOKVTL 120
Qy 121 NVHSEYPSWEQPMPTPKWFLPFLFOMPAYASSSKFVTRVTPAQRGKVPVHSSP 180
Db 121 NVHSEYPSWEQPMPTPKWFLPFLFOMPAYASSSKFVTRVTPAQRGKVPVHSSP 180
Qy 181 TQITHRPRVSRASSVAGDKPRTFLPSTTASKISALPGLLQKQTESYNHHTRLRORALD 240
Db 181 TQITHRPRVSRASSVAGDKPRTFLPSTTASKISALPGLLQKQTESYNHHTRLRORALD 240
Qy 241 YGSGREGQGFHILPTILGLFLALGLVYKRAVERKALSRRLAVRMALESSQ 300
Db 241 YGSGREGQGFHILPTILGLFLALGLVYKRAVERKALSRRLAVRMALESSQ 300
Qy 301 RPRGSPRPRSONNIYSACPRRARGADAAGTGEAPVPGCAPLPPAPLOVSESPLHAPSL 360
Db 301 RPRGSPRPRSONNIYSACPRRARGADAAGTGEAPVPGCAPLPPAPLOVSESPLHAPSL 360
Qy 361 KTSCEVSVLYHOPAMMEDSDSDYINVPA 390
Db 361 KTSCEVSVLYHOPAMMEDSDSDYINVPA 390

RESULT 2
Q9D8T1 ID Q9D8T1 PRELIMINARY; PRT; 422 AA.
AC Q9D8T1
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE 1810037B05RIK PROTEIN.
GN 1810037B05RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa K., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher M., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann C., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kotsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK007714; BAB25207.1;
DR MGI; MGI:1916419; 1810037B05RIK.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00410; IG_Like; 1.
SQ SEQUENCE 422 AA; 47532 MW; 2597083A50AD8E6E CRC64;

Query Match 56.3%; Score 1157; DB 11; Length 422;
Best Local Similarity 57.5%; Pred. No. 5.8e-95;
Matches 234; Conservative 46; Mismatches 107; Indels 20; Gaps 6;
Qy 1 MDRWLPYFLPVSGALRILPEVKVEGELGSGVTIKCPLEMHVRIYLCREMGSGTCCT 60
Db 1 MDRWLPYFLPVSGALRILPEVKVEGELGSGVTIKCPLEMHVRIYLCREMGSGTCCT 60
Qy 61 VSTTNFKAIEYKGRVTLKQYPRKNIPLVEVTOLTESDSGVYACGAGMNTDRGKTOKVTL 120
Db 61 VSTTNFKAIEYKGRVTLKQYPRKNIPLVEVTOLTESDSGVYACGAGMNTDRGKTOKVTL 120
Qy 121 NVHSEYPSWEQPMPTPKWFLPFLFOMPAYASSSKFVTRVTPAQRGKVPVHSSP 174
Db 121 NVHSEYPSWEQPMPTPKWFLPFLFOMPAYASSSKFVTRVTPAQRGKVPVHSSP 179
Qy 175 VHSSTPTQITHRPRVSRASSVAGDKPRTFLPSTTASKISALEGLLQKQTESYNHHTRLH 234
Db 175 VHSSTPTQITHRPRVSRASSVAGDKPRTFLPSTTASKISALEGLLQKQTESYNHHTRLH 238
Qy 235 RORALDYGSQSGREGQGFHILPTILGLFLALGLVYKRAVERKALSRRLAVRMALESSQ 287
Db 235 RORALDYGSQSGREGQGFHILPTILGLFLALGLVYKRAVERKALSRRLAVRMALESSQ 298
Qy 288 RLAVRMALESSQ-----QPRGSPRPRSONNIYSACPRRARGADAAGTGEAPVPGCAPL 342
Db 288 RLAVRMALESSQ-----QPRGSPRPRSONNIYSACPRRARGADAAGTGEAPVPGCAPL 358
Qy 343 PPAPLOVSESPLHAPSLKTSCEVSVLYHOPAMMEDSDSDYINVP 389
Db 343 PPAPLOVSESPLHAPSLKTSCEVSVLYHOPAMMEDSDSDYINVP 405
Qy 405 SPASQVLEAPWPHPTPSLKMCEVSVLYHOPAMMEDSDSDYINVP 405

RESULT 3
Q96SA2 ID Q96SA2 PRELIMINARY; PRT; 534 AA.
AC Q96SA2
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE FKSG87 PROTEIN.
GN FKSG87.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang Y.-G., Gong L.;
RT "Molecular cloning and characterization of FKSG87, a novel gene
located on human chromosome 1.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF34295; AAK39522.1;
SQ SEQUENCE 534 AA; 56748 MW; 6EF8050E412AF91C CRC64;

Query Match 9.1%; Score 186; DB 4; Length 534;
Best Local Similarity 22.2%; Pred. No. 3.5e-08;
Matches 98; Conservative 57; Mismatches 151; Indels 136; Gaps 20;
Qy 3 RMLWLPYFLPVSGALRILPEVKVEGELGSGVTIKCP 38
Db 45 RMLWE-GSLFRTSLRMTGLTSPSPCLWRESSEFAAPNSLKSRLVSGEPGAVTIQCH 103
Qy 39 LP-----EMHVRIYLCREMGSGTCGVSTTNFKAIEYKGRVTLKQYPRKNIPLVEVTOL 94
Db 104 YAPSVNRQRKYWCRLPPLPWCIVSTNYTHRRYDRVALTFPQGLVVRVUSQL 163
Qy 95 TESDSGVYACGAGMNTDRGKTOKVTLNVHSEYPSWEQPMPTPKWFLPFLFOMPAY 153
Db 164 SPDDICGICGIG-----SENNMLFLSMNLTLISAGPA---STLPTATPAAGEL---TMSY 213
Qy 154 ASSSKFVTRVTPAQRGKVPVHSSPTQITHRPRVSRASSVAGDK-----PRTFPLSTTAS 211


```
Db 214 GPASPVANRWTP-----GTTQTLQOQTANDTVASTPCTSKTTASAEGRTPGATREAPPGT 269
Qy 212 KISALEGLLK-----POTPSYNHHTRLHRQALDYGSQSGREGOGFHLIPTLGLFLULA 266
Db 270 G-SWAGSVKAPAPIPEPPSPKSRMSNTTIEGVWEGTRSS----- 308
Qy 267 LGLLVVKRA---VERKALSRARR-----LAVRMRALESSOR----- 301
Db 309 ----VTNRARASKDRREMTTKADRPREDIEGVRI-ALDAARKVLGTGPPALVSETLAW 363
Qy 302 ---PRGSP--RPRSONNIYSACPRRARGADAAGTGEAPV----- 335
Db 364 EILQATPVSKQSQSGSIGETTP--AAGWTLTGTPAADVWITSMEASGESGAAGDLDA 421
Qy 336 ---PGGAPLPPAPLOVSESPW 354
Db 422 TGDGRGPQATLSQTP---AVGPW 440

RESULT 4
Q9N2H7
ID Q9N2H7 PRELIMINARY; PRT; 758 AA.
AC Q9N2H7;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE POLY-IG RECEPTOR PRECURSOR.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP Sone T., Kumura H.;
RT "Porcine mammary gland cDNA clone, similar to poly-Ig receptor.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DDJB databases.
DR EMBL; AB032195; BAA84283.2; -.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF00047; Ig; 5.
DR SMART; SM00409; IG; 3.
DR SMART; SM00410; IG_Like; 2.
KW Signal; Receptor.
FT SIGNAL 1 18 POTENTIAL.
SQ SEQUENCE 758 AA; 83154 MW; D5BEBIA8B082D247 CRC64;

Query Match 9.0%; Score 184; DB 6; Length 758;
Best Local Similarity 34.2%; Pred. No. 8.3e-08;
Matches 41; Conservative 22; Mismatches 41; Indels 16; Gaps 4;

Qy 17 LRILPEVKGEL-----GGSVTTKCLPLP-----EMHVRIYLCREMAGSGTCGTVV 62
Db 10 LAIFPVVSMKSPFGPDQVSSVSGSSVIRCYYPATSVNRHQRKYLWCR-IGAKGRCTILI 68
Qy 63 STTNFIKAEYKGRVTLKQYPRKNLFLVEVTLTQTESDGVYACGAGMNTDRGKTQKVTLANV 122
Db 69 SSEGYSISKYGRANLUNTFNPENGTFVMDIGHLTRGDSGLYKCGLGISS-RLGSFVDSLEV 127

RESULT 5
Q920L8
ID Q920L8 PRELIMINARY; PRT; 455 AA.
AC Q920L8;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE FCA/M RECEPTOR (FRAGMENT).
GN FCAMR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
```

```
RN [1]
RP SEQUENCE FROM N.A.
RA Shimizu Y., Honda S., Yotsumoto K., Tahara-Hanaoka S., Eyre H.J.,
RA Sutherland G.D., Endo Y., Shibuya K., Koyama A., Nakauchi H.,
RA Shibuya A.;
RT "Fca/m receptor is a single gene-family member closely related to
RT polymeric immunoglobulin receptor on chromosome 1.";
RL Immunogenetics 0:0-0(2001).
DR EMBL; AB071978; BAB71750.1; -.
KW Receptor.
FT NON_TER 1
SQ SEQUENCE 455 AA; 48810 MW; 628913C33A4AC365 CRC64;

Query Match 8.5%; Score 175; DB 11; Length 455;
Best Local Similarity 23.5%; Pred. No. 2.7e-07;
Matches 89; Conservative 39; Mismatches 136; Indels 114; Gaps 14;

Qy 25 VEGELGGSVTIKCLPLP-----EMHVRIYLCREMAGSGTCGTVVSTNFIKAEYKGRVTLKQ 80
Db 10 VTGNTGGAVTIHCHYAPSSVNRHQRKYLWCRIGSLPLWICHTVVTNQYTHPDYRGRAALTD 69
Qy 81 YPRKNLFLVEVTLTQTESDGVYACGAGMNTDRGKTQKVTLANVHSEYEPS---WEEQPMPE 137
Db 70 VPQSGLFVVRLLRLSLGDLVGLYRCGIG---DRNDMLFVSVMLTVSAGPSNTTYAAAPASS 126
Qy 138 TPKWFHLPYLPQMPAYASSSKFVTRVTTPAQRGKVPVPHHSSPTTQIHRPRVSRASSVA 197
Db 127 EP-----TTASPCAASS-----ACNG-----WTSGVTQILEG-----S 154
Qy 198 GDKPRTFLPSTASKISALEGLLKPPQPSYNNHHTLRQALDYGSQSGREGOGFHLIP 257
Db 155 GSEMDRTAPTGTGTSK-----TTSSANGROTLTARTVVLGTGSRREGS----- 197
Qy 258 TILGLFLALLGLLVKRAVERRKALSRARLAVRMR-----ALESSQRPGR 304
Db 198 -----IRAAVPTPEGPSPKSRMSSTTQCVLWLNTRNSVTPSVTTSEGRQ 243
Qy 305 SPRPRSONNIYSACPRRARGADAAGTGEAPVPGCAPLPPAPL-----QVSE 351
Db 244 GTTPETDG-----PRDE--TDVRSPEAPRKTTGTPRSALISEHVHTWETLQDKTEVSK 295
Qy 352 SPWLH-----APSLKT 362
Db 296 QOMLHSLLEELSPAPSAQT 313

RESULT 6
Q9EQT7
ID Q9EQT7 PRELIMINARY; PRT; 535 AA.
AC Q9EQT7;
DT 01-MAR-2001 (TremBLrel. 16, Created)
DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE FCA/M RECEPTOR.
GN FCAMR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21170225; Pubmed=11062505;
RA Shibuya A., Sakamoto N., Shimizu Y., Shibuya K., Osawa M.,
RA Hiroyama T., Eyre H.J., Sutherland G.R., Endo Y., Fujita T.,
RA Miyabayashi T., Sakano S., Tsuji T., Nakayama E., Phillips J.H.,
RA Lanier L.L., Nakauchi H.;
RT "Fca/m receptor mediates endocytosis of IgM-coated microbe.";
RL Nat. Immunol. 1:441-446(2000).
DR EMBL; AB048834; BAB17312.1; -.
DR MGD; MGI:1927803; Fcamr.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003600; Ig_Like.
DR InterPro; IPR003006; Ig_MHC.
```

DR InterPro: IPR000508; Peptidase_S26.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00409; IG; 1.
DR SMART: SM00410; IG.Like; 1.
DR PROSITE: PS00501; SPASE_1_1; UNKNOWN_1.
DR Receptor.
SQ SEQUENCE 535 AA; 57696 MW; B275B77C70151C75 CRC64;

Query Match 8.5%; Score 175; DB 11; Length 535;
Best Local Similarity 23.5%; Pred. No. 3.3e-07;
Matches 89; Conservative 39; Mismatches 136; Indels 114; Gaps 14;

QY 25 VEGELGGVTTIKCLP-----EMHVRIVLCREMAGSGTCCTVVSTNTIKAEYKGRVTLKQ 80
DB 90 VTGNTGGAVTHCHYAPSSVNRHQKRWCRGLGSLWICHVTNVSTNQTHPDYGRALTD 149
QY 81 YPKNLFVLEVTQLTSDSGVYACGAGNMTDRGKTQKVTLVNHSEYEPS---WEEQPMPE 137
DB 150 VFQSLGFVYRLRLSLGDLVRCIG---DRNDMLFFSVNLTVSAGFSNTVAAAPASS 206
QY 138 TPKWPHLPYLFQMPAYASSSKFVTRVTPAQRGKVPVPHHSSPTQITHRPRVSRASSVA 197
DB 207 EP-----TTASPGAASS-----ACNG-----WTSGVTQILEG-----S 234
QY 198 GDKPTFLPSTTASKISALEGLLKPPQPSYNHHTLHQRALDYGSGSGREGQGFHILIP 257
DB 235 GSEWDRTPATPTGTSK-----TTSSANGROTTLRTARTVVLGTGSGREGS----- 277
QY 258 TILGLFLLALLGLVWKRVAERKALSRARRLAVRMK-----ALESSQPRG 304
DB 278 -----IRAAVPTPEGSPKSRSMSTTQGVWLNTRNVTSPVITSEGRQ 323
QY 305 SPRPSQNNIYACPRRAGDAAGTGEAPVCPGAPLPPAPL-----QVSE 351
DB 324 GTTPETDG-----PRDE--TDVRVSPAPRKTGTTPRSALISERHVTWETLQDKTEVSK 375
QY 352 SPWLH-----APSLKT 362
DB 376 QQLHLSLEELSPAPSAQT 393

RESULT 7
054947 ID O54947 PRELIMINARY; PRT; 307 AA.
AC O54947;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE KIDNEY INJURY MOLECULE-1 PRECURSOR (KIM-1).
GN KIM-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE=KIDNEY;
RX MEDLINE=98129827; PubMed=9461608;
RA Ichimura T., Bonventre J.V., Bailly V., Wei H., Hession C.A.,
RA Cate R.L., Sanicola M.;
RT *Kidney injury molecule-1 (KIM-1), a putative epithelial cell adhesion
RT molecule containing a novel immunoglobulin domain, is up-regulated in
RT renal cells after injury.";
RL J. Biol. Chem. 273:4135-4142(1998).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).
CC -!- TISSUE SPECIFICITY: EXPRESSED AT LOW LEVELS IN LIVER, SPLEEN AND
CC NORMAL KIDNEY. LEVELS INCREASE IN THE POSTISCHEMIC KIDNEY WITH
CC EXPRESSION FOUND IN REGENERATING PROXIMAL TUBULE EPITHELIAL CELLS.
CC -!- INDUCTION: IN RENAL CELLS, AFTER INJURY.
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS
CC ONE V-LIKE DOMAIN.
DR EMBL; AF035963; AAC53546.1; -.

DR InterPro: IPR003599; Ig.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00409; IG; 1.
DR Transmembrane; Cell adhesion; Glycoprotein; Signal.
KW SIGNAL 1 21
FT CHAIN 22 307
FT DOMAIN 22 235
FT TRANSMEM 236 256
FT DOMAIN 257 307
FT DOMAIN 22 130
FT DOMAIN 33 37
FT DOMAIN 131 201
FT DOMAIN 131 234
FT DOMAIN 138 142
FT REPEAT 143 147
FT REPEAT 148 152
FT DISULFID 37 108
FT CARBOHYD 206 206
FT CARBOHYD 218 218
SQ SEQUENCE 307 AA; 33963 MW; 736D1DD1F1549760 CRC64;

Query Match 7.8%; Score 161; DB 11; Length 307;
Best Local Similarity 25.1%; Pred. No. 2.9e-06;
Matches 89; Conservative 52; Mismatches 119; Indels 94; Gaps 20;

QY 13 VSGALRIIP-----EVKVEGELGGSVTIKCLPDMHVRIYLC--REMAGSGTCGTVVST 64
DB 8 ISGLLLLLGVSDVSEV-VKGVGHVPVTPCTYSTRGGITTCWGRGQCPSYSCONILIW 66
QY 65 TFIKAEYK--GRVTLKQYPRKNLFVLEVTQLTSDSGVYACGA--GMNDRGKTQKVT 119
DB 67 TNGVTVYTRSGRYNLIKRISEGDVSLTIENSVDGLYCCRVETPGWFND----QKMT 122
QY 120 LNVHSEYFSEWEEQPMPEPKWPHLPYLFQMPAYASSSKFVTRVTPAQRGKVPV---V 175
DB 123 FSL--EVKPEITSP--PTRP-----TTTPRTTPTTISTRTHVTPSTRV 165
QY 176 HSSPT---TQIHRPRVSRASSVAGDKPRTFLPSTTASKISALEGLLKPPQPSY----- 227
DB 166 STSPTPEQTQ--THRPEIT-----TFYAHETAEVT-----ETPSYTPADW 205
QY 228 -----NHETRLHQRALDYGSGSGREGQGFHILIPILGLFLLALLGLVWKR 275
DB 206 NGVTYSEANNNHTVRIPLRK-----PQRNPTAGFYGVGKSVAAALLLLLASTVVVTRY 259
QY 276 VERKK---ALSRARRLAVRMRALESSQPRGSPRPSQNNIYACPRRAGD 326
DB 260 IIRKKMGSLSFVAFHVS-KSRALQNA---AIVHPRAEDNIY-ITEDSRGAE 307

RESULT 8
Q9UBK4 ID Q9UBK4 PRELIMINARY; PRT; 299 AA.
AC Q9UBK4;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE IRL1.
GN IRL1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPHOID;
RA O'Connor C.D.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPHOID;


```

Db 188 TSV---PTTSIPTTSVPVTTVTSTFVPPMPPLPRQNHPEVATSPSPQPAETHETTLQG 244
Qy 219 LKPO---TPSYNHHT-----RLHRQALDYGSQSGREGGPHILPTIL 260
Db 245 AIRREPTSSPLYSYTTDGDNDVTESDGLWNNNOTQLFLEHSLLTANTTKGIYAGV-CIS 303
Qy 261 GLFLALLGLVVKRAVERRKALSRARRL-AVRMRALSSORPRGSPRPRSONNIY 315
Db 304 VLVLALLGLVIAKKYFFKKEVOQLSVFSLSLOKALONAV-----EKEVQAEADNIY 355

```

Search completed: October 28, 2002, 17:32:33
 Job time : 39.121 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	185	14.8	769	1	QRRTGS	secretory component
2	180.5	14.4	764	1	ORRHGS	secretory component
3	179.5	14.3	757	1	S48841	secretory component
4	179	14.3	757	2	I45956	polymorphic immunoglobulin
5	166.5	13.3	773	1	QRBBG	secretory component
6	100.5	8.0	504	2	A56205	transcription factor
7	100	8.0	224	2	I37243	CMRF-35 antigen -
8	96.5	7.7	847	2	JH0371	B-cell adhesion protein
9	93	7.4	346	2	S17475	STE50 protein - yeast
10	92.5	7.4	504	2	I49257	NP289 - mouse
11	92	7.3	355	1	LKCH	proteoglycan link
12	92	7.3	1906	1	S68235	myosin-light-chain
13	91.5	7.3	502	2	C56205	transcription factor
14	90.5	7.2	534	2	T22154	hypothetical protein
15	90.5	7.2	556	2	I45066	steroid hormone receptor
16	90.5	7.2	1839	1	RRPEEM	genome polyprotein
17	89.5	7.1	1415	1	IEDEGA	immediate-early protein
18	89	7.1	1241	2	S01827	period clock protein
19	88.5	7.1	328	2	JQ0985	hydroxyproline-rich
20	88.5	7.1	1021	2	T42634	connectin/titin
21	87	6.9	249	2	S69340	Ig heavy chain VH1
22	87	6.9	1070	2	T34385	hypothetical protein
23	86.5	6.9	228	2	JG7761	dendritic cell protein
24	86.5	6.9	456	2	T45610	proanthranilate N-hydroxyproline-rich
25	86	6.9	303	2	S28264	nuclear receptor corepressor
26	86	6.9	2453	2	S60254	Ig gamma chain (5'
27	85.5	6.8	342	2	A46529	hypothetical protein
28	85.5	6.8	496	2	T01564	alpha-globin transcript
29	85.5	6.8	501	2	A42030	trans

secretory component precursor [validated] - human
 N;Alternate names: poly-Ig receptor; polymeric immunoglobulin receptor
 N;Contains: free secretory component; transmembrane secretory component
 C;Species: Homo sapiens (man)
 C;Date: 28-Aug-1985 #sequence_revision 23-Aug-1996 #text_change 08-Dec-2000
 C;Accession: A46537; A55284; J38115; A32263; S38978; S13453; A02112
 R;Krajci, P.; Kvate, D.; Tasken, K.; Brandtzaeg, P.
 Eur. J. Immunol. 22, 2309-2315, 1992
 A;Title: Molecular cloning and exon-intron mapping of the gene encoding human transmembrane secretory component
 A;Reference number: A46537; MUID:92387236
 A;Accession: A46537
 A;Status: not compared with conceptual translation
 A;Molecule type: DNA
 A;Residues: 1-764 <KRA>
 A;Cross-references: GB:S43449; NID:q255097; PIDN:AA23176.1; PID:q255098
 A;Experimental source: leukocytes
 A;Note: sequence extracted from NCBI backbone (NCBIP:113253)
 R;Krajci, P.; Grzeschik, K.H.; Geurts van Kessel, A.H.; Olaisen, B.; Brandtzaeg, P.
 Hum. Genet. 87, 642-648, 1991
 A;Title: The human transmembrane secretory component (poly-Ig receptor): molecular cloning and exon-intron mapping
 A;Reference number: A55284; MUID:92039621
 A;Accession: A55284
 A;Molecule type: mRNA
 A;Residues: 1-764 <KR2>
 A;Cross-references: GB:S62403; NID:q238235; PIDN:AA220203.1; PID:q238236
 A;Experimental source: colonic adenocarcinoma cell line
 A;Note: sequence extracted from NCBI backbone (NCBIN:62403, NCBIP:62408)
 R;Piskurich, J.F.; France, J.A.; Tamer, C.M.; Willmer, C.A.; Kaetzel, D.M.
 Mol. Immunol. 30, 413-421, 1993
 A;Title: Interferon-gamma induces polymeric immunoglobulin receptor mRNA in human intest
 A;Reference number: I38115; MUID:93205018
 A;Accession: I38115
 A;Molecule type: mRNA
 A;Residues: 1-764 <RES>
 A;Cross-references: EMBL:X73079; NID:q456345; PIDN:CAA51532.1; PID:q456346
 A;Note: submitted to the EMBL/GenBank/DBJ databases by J.F. Piskurich, February 1994
 R;Krajci, P.; Solberg, R.; Sandberg, M.; Oyen, O.; Jahnsen, T.; Brandtzaeg, P.
 Biochem. Biophys. Res. Commun. 158, 783-789, 1989
 A;Title: Molecular cloning of the human transmembrane secretory component (poly-Ig recept
 A;Reference number: A32263; MUID:89149795
 A;Accession: A32263
 A;Molecule type: mRNA
 A;Residues: 72-764 <KR3>
 A;Cross-references: GB:M24559; NID:q514365; PIDN:AAA36102.1; PID:q514366
 R;Fallgreen-Gebauer, E.; Gebauer, W.; Bastian, A.; Kratzin, H.D.; Eiffert, H.; Zimmermann
 Biol. Chem. Hoppe-Seyler 374, 1023-1028, 1993
 A;Title: The covalent linkage of secretory component to IgA. Structure of sigA.
 A;Reference number: S38978; MUID:94121784
 A;Accession: S38978
 A;Molecule type: protein
 A;Residues: 478-488; 517-526; 543-545 <FAL>
 A;Note: disulfide bonds for unbound and IgA-bound forms
 R;Eiffert, H.; Quentin, E.; Wiederhold, M.; Hillemeir, S.; Decker, J.; Weber, M.; Hilsch
 Biol. Chem. Hoppe-Seyler 372, 119-128, 1991
 A;Title: Determination of the molecular structure of the human free secretory component.
 A;Reference number: S13453; MUID:91315750
 A;Accession: S13453
 A;Molecule type: protein
 A;Residues: 19-135, 'Q', 137-157, 'D', 159-207, 'DE', 210-228, 230-233, 'N', 235-240, 'Q', 242-261,
 R;Eiffert, H.; Quentin, E.; Decker, J.; Hillemeir, S.; Hufschmidt, M.; Klingmuller, D.;
 Hoppe-Seyler's Z. Physiol. Chem. 365, 1489-1495, 1984
 A;Title: The primary structure of the human free secretory component and the arrangement
 A;Reference number: A02112; MUID:85128981
 A;Accession: A02112
 A;Molecule type: protein
 A;Residues: 19-157, 'D', 159-207, 'DE', 210-228, 230-233, 'N', 235-240, 'Q', 242-261, 'Q', 263-279,
 A;Note: paper in German with English abstract
 C;Comment: As a 100K transmembrane receptor for polymeric immunoglobulin, secretory com
 ylated, forms interchain disulfide bonds, undergoes proteolysis and transcytosis. Free s
 C;Genetics:
 A;Gene: GDB:PTGR
 A;Cross-references: GDB:I20290; OMIM:173880
 A;Map position: lq31-lq41

A;Introns: 15/1; 130/1; 349/1; 460/1; 569/1; 629/2; 670/1; 714/1; 733/3
 A;Note: the first intron occurs before the initiator codon
 C;Complex: monomeric as a transmembrane receptor or free in mucosal secretions; heter
 amers; hetero-22-mer composed of one chain of secretory component, one chain of immun
 C;Superfamily: secretory component; immunoglobulin homology
 C;Keywords: duplication; glycoprotein; immunoglobulin receptor; phosphoprotein; trans
 F;1-18/Domain: signal sequence #status predicted <SIG>
 F;19-764/Product: transmembrane secretory component #status predicted <NATM>
 F;19-577/Product: free secretory component #status experimental <MATF>
 F;33-112/Domain: immunoglobulin homology <IM1>
 F;145-222/Domain: immunoglobulin homology <IM2>
 F;250-327/Domain: immunoglobulin homology <IM3>
 F;364-443/Domain: immunoglobulin homology <IM4>
 F;475-546/Domain: immunoglobulin homology <IM5>
 F;639-661/Domain: transmembrane #status predicted <TMM>
 F;662-764/Domain: intracellular #status predicted <INT>
 F;40-110,56-64,152-220,257-325,271-279,371-441,385-395,482-544,496-503/Disulfide bond
 F;83,90,135,186,421,469,499/Binding site: carbohydrate (Asn) (covalent) #status exper
 F;486-520/Disulfide bonds: (in Ig-unbound form) #status experimental
 F;486/Disulfide bonds: interchain (to IgA alpha-1 chain-192) #status experimental
 F;520/Binding site: cysteine (Cys) (covalent) (in Ig-bound form) #status experimental
 F;577-578/Cleavage site: Lys-Ala (unidentified proteinase) #status experimental
 F;673/Binding site: phosphate (Ser) (covalent) #status predicted
 Query Match 14.4%; Score 180.5; DB 1; Length 764;
 Best Local Similarity 38.7%; Pred. No. 9.1e-07;
 Matches 41; Conservative 19; Mismatches 39; Indels 7; Gaps 4;
 QY 4 PEVKVEGLGSGVTIKCPPLP----EMHVIYLICREMGSGTGVVSTTFNFIKAEYKGRV 59
 Db 25 PE-EVNSVEGNSVITCYPPYSVNRHTRKYWCROGARGG-CITLISSEGVSSKYAGRA 82
 QY 60 TLKQPKRNFLVEVTVLTESDSGVYACGAGMTDRGKTQKTVLNV 105
 Db 83 NLTFNPENGTFFVNIQAQSDDSGRYKGLGINS-EGUSFDVSLV 127
 RESULT 3
 S48841
 secretory component precursor - bovine
 N;Alternate names: poly-Ig receptor; polymeric immunoglobulin receptor
 N;Contains: free secretory component; transmembrane secretory component
 C;Species: Bos primigenius taurus (cattle)
 C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C;Accession: S48841
 R;Vermeer, H.; Warmerdam, G.W.; de Boer, H.A.; Verbeet, M.P.H.
 submitted to the EMBL Data Library, September 1994
 A;Description: The cloning, tissue specific expression and interspecies sequence comp
 A;Reference number: S48841
 A;Accession: S48841
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-757 <VER>
 A;Cross-references: EMBL:X81371; NID:q563340; PIDN:CAA57136.1; PID:q563341
 C;Superfamily: secretory component; immunoglobulin homology
 C;Keywords: duplication; glycoprotein; immunoglobulin receptor; phosphoprotein; trans
 F;1-18/Domain: signal sequence #status predicted <SIG>
 F;19-572/Product: transmembrane secretory component #status predicted <NATM>
 F;19-577/Product: free secretory component #status predicted <MATF>
 F;33-112/Domain: immunoglobulin homology <IM1>
 F;145-222/Domain: immunoglobulin homology <IM2>
 F;250-328/Domain: immunoglobulin homology <IM3>
 F;365-444/Domain: immunoglobulin homology <IM4>
 F;476-547/Domain: immunoglobulin homology <IM5>
 F;631-653/Domain: transmembrane #status predicted <TMM>
 F;654-757/Domain: intracellular #status predicted <INT>
 F;40-110,56-64,152-220,166-173,257-324,271-279,370-440,384-394,481-543,495-502/Disulf
 F;83,420,468/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F;665/Binding site: phosphate (Ser) (covalent) #status predicted
 Query Match 14.3%; Score 179.5; DB 1; Length 757;
 Best Local Similarity 37.6%; Pred. No. 1.1e-06;
 Matches 41; Conservative 19; Mismatches 42; Indels 7; Gaps 4;

R:Jackson, D.G.; Hart, D.N.; Starling, G.; Bell, J.I.

Eur. J. Immunol. 22, 1157-1163, 1992

A:Title: Molecular cloning of a novel member of the immunoglobulin gene superfamily hom

A:Reference number: 137243; MUID:92249405

A:Accession: J137243

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-224 <RES>

A:Cross-references: EMBL:X66171; NID:g396169; PIDN:CAA46948.1; PID:g396170

C:Genetics:

A:Gene: CMK35

Query Match 8.0%; Score 100; DB 2; Length 224;
Best Local Similarity 31.3%; Pred. No. 0.7;
Matches 26; Conservative 15; Mismatches 38; Indels 4; Gaps 3;

QY 8 VGEIGGGVTKCPLEPMH--VRIYLREMGSGTGTGVSTTNPIKAETKGRVTLKQYP 65

Db 31 VAGPVGGSLSVQCYREKERTLNKFWCHP-PQILRCDKIVETKG-SACKRNGRVSIROSP 88

QY 66 RKNLFLVEVTOLTESDGVYACG 88

Db 89 ANLSFTVTLNLTEDAGTYWCG 111

RESULT 8

JH0371

N:cell adhesion protein CD22 beta splice form precursor - human

A:Alternate names: B-cell membrane protein CD22

C:Species: Homo sapiens (man)

C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 31-Jan-2000

C:Accession: JH0371; I56171

R:Wilson, G.L.; Fox, C.H.; Fauci, A.S.; Kehrl, J.H.

J. Exp. Med. 173, 137-146, 1991

A:Title: cDNA cloning of the B cell membrane protein CD22: a mediator of B-B cell intera

A:Reference number: JH0371; MUID:91086838

A:Accession: JH0371

A:Molecule type: mRNA

A:Residues: 1-847 <WTLI>

A:Cross-references: GB:X59350; NID:g36090; PIDN:CAA42006.1; PID:g36091

A:Experimental source: B lymphocyte

A>Note: the authors translated the codon AAT for residue 358 as Met

R:Wilson, G.L.; Najfeld, V.; Kozlow, E.; Menniger, J.; Ward, D.; Kehrl, J.H.

J. Immunol. 150, 5013-5024, 1993

A:Title: Genomic structure and chromosomal mapping of the human CD22 gene.

A:Reference number: I56171; MUID:93267103

A:Accession: I56171

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 121-269, 'T', 271-473, 'K', 475-614, 'R', 616-638, 'Y', 640-711, 777-847 <WIL2>

A:Cross-references: GB:S61375; NID:g385980; PIDN:AAC18956.1; PID:g3184492

C:Genetics:

A:Gene: GDB:CD22

A:Cross-references: GDB:127545; OMIM:107266

A:Map position: 19q13.1-19q13.1

A:Introns: 138/1; 240/1; 329/1; 417/1; 503/1; 591/1; 679/1; 711/2; 804/3

C:Superfamily: immunoglobulin homology

C:Keywords: alternative splicing; B-cell; cell adhesion; dimer; glycoprotein; phosphop

F:1-19/Domain: signal sequence #status predicted <Sig>

F:20-847/Product: B lymphocyte cell adhesion protein #status predicted <NAT>

F:346-398/Domain: immunoglobulin homology <IMM1>

F:609-661/Domain: immunoglobulin homology <IMM2>

F:688-706/Domain: transmembrane #status predicted <TRA>

F:67,101,112,135,164,231,363,445,448,479,574,634/Binding site: carbohydrate (Asn) (coval

F:764,789/Binding site: phosphate (Thr) (covalent) #status predicted

Query Match 7.7%; Score 96.5; DB 2; Length 847;

Best Local Similarity 20.7%; Pred. No. 6.2;

Matches 42; Conservative 31; Mismatches 69; Indels 61; Gaps 8;

QY 1 RILPEKVEGELGGSVTIKCPLEPMHVRIYLREMGSGTGTGVSTTNPIKAETKGRVT 60

Db 247 KVTPESDAIVRE-GDSVTMTCT-----EVSSSNPEYTTVSM-----LKDGTS 285

QY 61 LKQYPRKNLFLVEVTOLTESDGVYACGAGMNTDRGKTQKVTLNHVSEYSPSWEE----- 115

Db 286 LK---KQNTFTLNREVTKDQSGKYCCQVSDVGPGRSEEVFLQVYAPESPSTVQILHSP 342

QY 116 -----QPMPEYPKWF-----HLPYLFQMPAYASSSKFVT-R 145

Db 343 AVESQVEFLCMLANPLPTNTVTHNCKEMOGRTEEKVHPKT--LPWHAGTYSYSCVAEN 400

QY 146 VTTPAQRGKVPVPHVHSSPTTQIT 168

Db 401 ILGTGORGPGCAELDVQYPPPKVT 423

RESULT 9

SL17475

STE50 protein - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein YCL032w; protein YCL185

C:Species: Saccharomyces cerevisiae

C>Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 21-Jul-2000

C:Accession: SL17475; S30158; S19360

R:Ramezani Rad, M.; Luetzenkirchen, K.; Xu, G.; Kleinhans, U.; Hollenberg, C.P.

Yeast 7, 533-538, 1991

A:Title: The complete sequence of a 11,953 bp fragment from CIG on chromosome III enc

A:Reference number: SL17471; MUID:91377317

A:Accession: SL17475

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-346 <RAM>

A:Cross-references: EMBL:X59720; NID:g1907116; PIDN:CAA42384.1; PID:g5332

R:Ramezani Rad, M.; Xu, G.; Hollenberg, C.P.

Mol. Gen. Genet. 236, 145-154, 1992

A:Title: STE50, a novel gene required for activation of conjugation at an early step

A:Reference number: S30158; MUID:93156679

A:Accession: S30158

A:Molecule type: DNA

A:Residues: 1-346 <RA2>

A:Cross-references: EMBL:Z11116; NID:g3312; PIDN:CAA77462.1; PID:g3313

R:Hollenberg, C.P.; Kleinhans, U.; Luetzenkirchen, K.; Ramezani Rad, M.; Xu, G.

submitted to the Protein Sequence Database, March 1992

A:Reference number: S19350

A:Molecule type: DNA

A:Residues: 1-346 <HOL>

A:Cross-references: EMBL:X59720; NID:g1907116; PID:e264426; PID:g5332; MIPS:YCL032w

C:Genetics:

A:Gene: SGD:STE50

A:Cross-references: SGD:S0000537; MIPS:YCL032w

A:Map position: 3L

C:Superfamily: SAM homology

F:32-101/Domain: SAM homology <SAM>

Query Match 7.4%; Score 93; DB 2; Length 346;

Best Local Similarity 22.2%; Pred. No. 4.2;

Matches 56; Conservative 43; Mismatches 101; Indels 52; Gaps 10;

QY 2 ILPEVKV-----EGELGGSVTIKCPLEPMHVRIYLREMGSGTGTGV-----STNLF 50

Db 69 LLPELCQDQCQDLCDGLDKAIFKILINKRDKSLKWKDKQEDMITVLKLYTTTSA 128

QY 51 IKASYGRVTLKQYPRKNLFLVEVTOLTESDGVYACG-----AGMNTDRGKTQKVTLN 104

Db 129 KLQEFQ-----SQYTLRMDVLDVDMKTSSTSSSPINTHGVTSTVPSSNNITLIPSSDGYSL 183

QY 105 -----VISEYEPWEQPMPEYKFWHLPLYLFQMPAYASSSKFVTRVTTTQAQGVKVP 157

Db 184 QTDYFDVFNHROSPPRRSPVT-----VFRQPSLSHKSLSL-----HSDSKNKVPQ 228

QY 158 VHHSSPTTQITTHRPVRSASSVAGDKPRFTLPTSTASKISALEGLLKPTQPSYNNHTRLH 217

Db 229 I-----STNOSHPSAVSTANT-PPGSPNEALQFLASKEDSCERILKNAMKRNLAODW 282

QY 218 RQRAL--DYGSQ 227

A:Residues: 1750-1906 <C02>
A:Cross-references: GB:M88283; NID:g211371; PIDN:AAA48647.1; PID:g211372
R.Yoshikawa, S.I.; Ikebe, M.
Arch. Biochem. Biophys. 299, 242-247, 1992
A:Title: Molecular cloning of the chicken gizzard telokin gene and cDNA.
A:Reference number: S28227; MUID:93073972
A:Accession: S28227
A:Molecule type: mRNA
A:Residues: 1750-1906 <YOS>
A:Cross-references: EMBL:M96655; NID:g212744; PIDN:AAA49083.1; PID:g212745
A:Accession: S78216
A:Molecule type: DNA
A:Residues: 1750-1906 <YOW>
A:Cross-references: EMBL:M96987
R.Olson, N.J.; Pearson, R.B.; Needleman, D.S.; Hurwitz, M.Y.; Kemp, B.E.; Means, A.R.
Proc. Natl. Acad. Sci. U.S.A. 87, 2284-2288, 1990
A:Title: Regulatory and structural motifs of chicken gizzard myosin light chain kinase.
A:Reference number: A35093; MUID:90192792
A:Accession: A35093
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 955-1438, 'Q', 1440-1906 <OLS>
A:Cross-references: GB:M31048; NID:g212660; PIDN:AAA49069.1; PID:g212661
R.Guerriero Jr., V.; Russo, M.A.; Olson, N.J.; Putkey, J.A.; Means, A.R.
Biochemistry 25, 8372-8381, 1986
A:Title: Domain organization of chicken gizzard myosin light chain kinase deduced from a
A:Reference number: A25810; MUID:87157587
A:Accession: A25810
A:Molecule type: mRNA
A:Residues: 1258-1438, 'Q', 1440-1906 <GUE>
C:Genetics: 1258-1438, 'Q', 1440-1906 <GUE>
A:Introns: 1735/3; 1779/1; 1819/1
C:Superfamily: myosin-light-chain kinase, nonmuscle; fibronectin type III repeat homolog
C:Keywords: alternative initiators; ATP; calmodulin binding; phosphoprotein; phosphotran
F:542-599/Domain: immunoglobulin homology <IMM1>
F:935-1906/Product: myosin-light-chain kinase, 108K, smooth muscle (from 5.5kb transcrip
F:1098-1158/Domain: immunoglobulin homology <IMM2>
F:1451-1708/Domain: protein kinase homology <KIN>
F:1459-1467/Region: protein kinase ATP-binding motif
F:1750-1906/Product: telokin (kinase-related protein KRP) (from 2.7 kb transcript) #stat
F:1808-1869/Domain: immunoglobulin homology <IMM3>

Query Match 7.3%; Score 92; DB 1; Length 1906;
Best Local Similarity 22.6%; Pred. No. 36;
Matches 49; Conservative 28; Mismatches 48; Indels 92; Gaps 12;

QY 12 LGGSVTIKCP---LPENHVRIVLCREMAGSGTGVVTTNFIKAETKGRVTLKQYPR-- 66
DB 19 LSPSVPAEAPFTLPPRNIRVQL-----GA-----TARFEGKV--RGYPEPQ 58

QY 67 -----KNLFLVEVTQLTESDSGVYACGAGMNTDRGKTQKVTILN 104
DB 59 ITWRNGHLEPGRHYVDHSIRGIFSLVKGQEGDSKTYCEA---ANDGVRQVIVE 115

QY 105 VHSEYEPSEWEEQMPETPKWFLHLYLFQMPAYASSKFFVTRVTPAQKGVPPVHHSSPT 164
DB 116 LTVE-----GNSLKKYSLP-----SSAK-----TPGRLSPVPE----- 145

QY 165 TQTHPRVSRASSVAGDKPRTFLPSTTAKISALEG 201
DB 146 ----HRP-----SINGEPPKF--ATKPNRVVREG 170

RESULT 13
C56205
Transcription factor LBP1c - human
N:Alternate names: alpha globulin transcription factor CP2
C:Species: Homo sapiens (man)
C:Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 01-Dec-2000
A:Accession: C56205; B53771
R.Yoon, J.B.; Li, G.; Roeder, R.G.
Mol. Cell. Biol. 14, 1776-1785, 1994
A:Title: Characterization of a family of related cellular transcription factors which co

A:Reference number: A56205; MUID:94158849
A:Accession: C56205
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-502 <YOO>
R.Shirra, M.K.; Zhu, Q.; Huang, H.C.; Pallas, D.; Hansen, U.
Mol. Cell. Biol. 14, 5076-5087, 1994
A:Title: One exon of the human LSF gene includes conserved regions involved in novel
A:Reference number: A53771; MUID:94309627
A:Accession: B53771
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-490, 'Q', 492-502 <SHI>
A:Cross-references: GB:U03494; NID:9476098; PIDN:AAA21324.1; PID:9476099
C:Keywords: alternative splicing; transcription factor

Query Match 7.3%; Score 91.5; DB 2; Length 502;
Best Local Similarity 23.5%; Pred. No. 8.6;
Matches 39; Conservative 21; Mismatches 57; Indels 49; Gaps 8;

QY 47 TTNPITAEYKGRVTLKQYPRKNLFLVEVTQLTESDSGVY-----AC-----GAG- 90
DB 193 STEETMRKHGGE---KGVF---FRVQIDTFKENENGEXTEHLHSACQIKVFKPKGADR 245

QY 91 -MNTDRGKTQKVTLVNHSSEYSEWEPQMPETPKWFLHLYLFQMPA---YASSSKFVTRV 146
DB 246 KQKTDREKMEKRTPEHEKQPSYETITLTCSPWPEITVNNVSPSGFNSSHSF---- 301

QY 147 TTPAQKGVPPVHHSSPTTQTHPRVSRASSVAGDKPRTFLPSTT 192
DB 302 --SLGEGNGSPNHQPEPPPVTD-----NLLPTTT 329

RESULT 14
T22154
Hypothetical protein F44A6.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
A:Accession: T22154
R.Sulston, J
submitted to the EMBL Data Library, August 1995
A:Reference number: Z19524
A:Accession: T22154
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-534 <WIL>
A:Cross-references: EMBL:Z50858; PIDN:CAA90722.1; GSPDB:GN00028; CESP:F44A6.2
A:Experimental source: clone F44A6
C:Genetics:
A:Gene: CESP:F44A6.2
A:Map position: X
A:Introns: 58/3; 123/1; 172/3; 220/1; 285/3; 464/1; 506/3
C:Superfamily: unassigned erba-related proteins; erba transforming protein homology

Query Match 7.2%; Score 90.5; DB 2; Length 534;
Best Local Similarity 25.1%; Pred. No. 11;
Matches 43; Conservative 19; Mismatches 78; Indels 31; Gaps 8;

QY 77 LTESDSGVYACGAGMNTDRGKTQKVTLVNHSSEYSEWEEQMPETPKWFLHLYLFQMPA 135
DB 12 LTNPDTPLSVCTSPYSPGSKTASI-----PSSEASKPECTNGQWHLPTGATYTV 52

QY 136 YASSSKF-VTRVTPAQKRGK----VPPVHHSSPTQI---THRPRVSRASSVAGDKPRTF 187
DB 63 DEFSSFEQFQNGSTAAQSGNANNYADPLSHRRYFNVNGVYNNHHQFYDTASQASVSSPAT- 121

QY 188 LPSTTAKISALEGLLKPQTPSYNNHHT--RLHRQALDYGSGSGREGGFH 236
DB 122 --SVTSS-----LSPDPSLSNGHTTQRIHCKAISFCCKVCGDKASGYH 162

RESULT 15
I45066

Search completed: October 28, 2002, 17:34:02
Job time : 15.371 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2002, 17:24:41 ; Search time 6.66129 Seconds
(without alignments)
1371.778 Million cell updates/sec

Title: US-09-135-238B-2_COPY_18_253

Perfect score: 1254

Sequence: 1 RILPEVKVEGLGGSTIKC.....HQRALDYGSGREGQGQH 236

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	194	15.5	771	1	P1GR_MOUSE
2	185	14.8	769	1	P1GR_RAT
3	180.5	14.4	764	1	P1GR_HUMAN
4	179	14.3	757	1	P1GR_BOVIN
5	166.5	13.3	773	1	P1GR_RABIT
6	120.5	9.6	102	1	P1GR_PIG
7	100	8.0	224	1	CM35_HUMAN
8	96.5	7.7	847	1	CD22_HUMAN
9	93	7.4	346	1	ST50_YEAST
10	93	7.4	2440	1	NCRL_HUMAN
11	92	7.3	355	1	PLK_CHICK
12	92	7.3	1906	1	KMLS_CHICK
13	91.5	7.3	872	1	FPI_MYTICO
14	90.5	7.2	534	1	NCRD_CAEEL
15	90.5	7.2	1839	1	POLR_EPMV
16	89.5	7.1	1415	1	ICP4_HSMVG
17	89	7.1	1241	1	PER_DROPS
18	88	7.0	835	1	AXNL_BRARE
19	86	6.9	2453	1	NCRL_MOUSE
20	85.5	6.8	1506	1	PK3G_MOUSE
21	85.5	6.8	3649	1	ACVS_NOCLA
22	85	6.8	283	1	EXTN_SORBI
23	84.5	6.7	704	1	MYB_MOUSE
24	84.5	6.7	739	1	FLOB_YEAST
25	84	6.7	275	1	IAP_GVCP
26	84	6.7	530	1	PVR2_MOUSE
27	83.5	6.7	146	1	HM2I_HUMAN
28	83.5	6.7	325	1	HM06_CAEEL
29	83.5	6.7	387	1	PIGM_RHOSO
30	83	6.6	5147	1	FAT1_DROME
31	82.5	6.6	358	1	KLFI_MOUSE
32	82	6.5	142	1	HV01_RAT
33	82	6.5	749	1	SM3B_HUMAN

34	81.5	6.5	636	1	KDPG_SCHPO	014019 schizosacch
35	81.5	6.5	912	1	PGCB_BOVIN	Q28062 bos taurus
36	81.5	6.5	1257	1	CAML_HUMAN	P32004 homo sapien
37	81.5	6.5	1259	1	CAML_RAT	Q05695 rattus norv
38	81.5	6.5	1448	1	PK3G_HUMAN	P03218 epstein-bar
39	81	6.5	248	1	VGL2_EBV	Q92854 homo sapien
40	81	6.5	862	1	SM4D_HUMAN	Q08575 mus musculu
41	80.5	6.4	532	1	EYA2_MOUSE	Q08575 mus musculu
42	80.5	6.4	938	1	EBN4_EBV	P03203 epstein-bar
43	80	6.4	518	1	TEX5_HUMAN	Q95993 homo sapien
44	80	6.4	571	1	FUB1_CAMJE	P56964 campylobact
45	80	6.4	688	1	YAQA_SCHPO	Q10109 schizosacch

ALIGNMENTS

RESULT 1
PIGR_MOUSE
ID PIGR_MOUSE STANDARD; PRT; 771 AA.
AC 070570;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Polymorphic-immunoglobulin receptor precursor (Poly-IG receptor) (PIGR)
DE [Contains: Secretory component].
GN PIGR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6 X CBA; TISSUE=Liver;
RX MEDLINE=95138517; PubMed=7836758;
RA Piskurich J.F., Blanchard M.H., Youngman K.R., France J.A.,
RA Kaetzel C.S.;
RT "Molecular cloning of the mouse polymeric Ig receptor. Functional
RT regions of the molecule are conserved among five mammalian species.";
RL J. Immunol. 154:1735-1747(1995).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ;
RX MEDLINE=98072444; PubMed=9409786;
RA Martin M.G., Gutierrez E.M., Lam J.T., Li T.W.H., Wang J.;
RT "Genomic cloning and structural analysis of the murine polymeric
RT receptor (pigr) gene and promoter region.";
RL Gene 201:189-197(1997).
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=129; TISSUE=Liver;
RA de Groot N., Vollebregt E., Lee S.H., Verbeet M.P., de Boer H.A.;
RT "Molecular cloning and exon-intron organization of the gene encoding
RT the murine polymeric immunoglobulin receptor.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC !- FUNCTION: THIS RECEPTOR BINDS POLYMERIC IGA AND IGM AT THE
CC BASOLATERAL SURFACE OF EPITHELIAL CELLS. THE COMPLEX IS THEN
CC TRANSPORTED ACROSS THE CELL TO BE SECRETED AT THE APICAL SURFACE.
CC DURING THIS PROCESS A CLEAVAGE OCCURS THAT SEPARATE THE
CC EXTRACELLULAR (KNOWN AS THE SECRETORY COMPONENT) FROM THE
CC TRANSMEMBRANE SEGMENT (BY SIMILARITY).
CC !- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. ALSO SECRETED.
CC !- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC !- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE V-TYPE DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).

```

NCBI_TaxID=9986;
[1]
SEQUENCE FROM N.A.
MEDLINE=84142246; PubMed=6322002;
RA Mostov K.E., Friedlander M., Blobel G.;
RT "The receptor for transepithelial transport of Iga and Igm contains
RT multiple immunoglobulin-like domains.";
RL Nature 308:37-43(1984).
[2]
SEQUENCE OF 87-114 AND 410-428.
RX MEDLINE=8828032; PubMed=3131339;
RA Frutiger S., Hughes G.J., Hanly W.C., Jaton J.-C.;
RT "Rabbit secretory components of different allotypes vary in their
RT carbohydrate content and their sites of N-linked glycosylation.";
RL J. Biol. Chem. 263:8120-8125(1988).
CC -!- FUNCTION: THIS RECEPTOR BINDS POLYMERIC IGA AND IGM AT THE
CC BASOLATERAL SURFACE OF EPITHELIAL CELLS. THE COMPLEX IS THEN
CC TRANSPORTED ACROSS THE CELL TO BE SECRETED AT THE APICAL SURFACE.
CC DURING THIS PROCESS A CLEAVAGE OCCURS THAT SEPARATE THE
CC EXTRACELLULAR (KNOWN AS THE SECRETORY COMPONENT) FROM THE
CC TRANSMEMBRANE SEGMENT.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. ALSO SECRETED.
CC -!- POLYMORPHISM: THE SEQUENCE SHOWN IS THAT OF ALLOTYPIC T62.
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE V-TYPE DOMAINS.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL: X00412; CAA25118.1; -.
PIR: A02111; ORR87.
DR A02111; A28077.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig_5.
DR SMART: SM00409; Ig_5.
DR SMART: SM00410; Ig_Like; 1.
KW Immunoglobulin domain; Repeat; Transmembrane; Glycoprotein; Signal;
FT Polymorphism.
FT CHAIN 1 18 POLYMERIC-IMMUNOGLOBULIN RECEPTOR.
FT CHAIN 19 773 SECRETORY COMPONENT.
FT DOMAIN 19 615 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 648 670 POTENTIAL.
FT DOMAIN 671 773 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 137 243 IG-LIKE V-TYPE DOMAIN 1.
FT DOMAIN 137 243 IG-LIKE V-TYPE DOMAIN 2.
FT DOMAIN 244 350 IG-LIKE V-TYPE DOMAIN 3.
FT DOMAIN 351 456 IG-LIKE V-TYPE DOMAIN 4.
FT DOMAIN 457 558 IG-LIKE V-TYPE DOMAIN 5.
FT DISULFID 46 115 POTENTIAL.
FT DISULFID 155 225 POTENTIAL.
FT DISULFID 260 324 POTENTIAL.
FT DISULFID 369 438 POTENTIAL.
FT DISULFID 478 538 POTENTIAL.
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .); IN ALLOTYPIC T61.
FT CARBOHYD 108 108 N-LINKED (GLCNAC. . .); IN ALLOTYPIC T62
(PARTIAL) AND T63.
FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .).
FT VARIANT 88 88 K -> N (IN ALLOTYPIC T61).
FT VARIANT 94 94 D -> E (IN ALLOTYPIC T61).
FT VARIANT 101 108 TVDQLTQN -> YLNRLSQS (IN ALLOTYPIC T61).
FT VARIANT 110 110 S -> T (IN ALLOTYPIC T63).
SQ SEQUENCE 773 AA; 83886 MW; DF2C4D2F1193C65 CRC64;
Query Match 13.3%; Score 166.5; DB 1; Length 773;
Best Local Similarity 42.7%; Pred. No. 2.3e-06;

```

```

Matches 44; Conservative 11; Mismatches 41; Indels 7; Gaps 4;
QY 13 GGSVTIKCPLP----EMHVRIYLCREMAGSGTCGTVVSTTNFKAEYKGRVTLKQYPRKN 68
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 39 GDSVSICTYYPTTSTVTRHSRKFWCRE--EESRCVTLAS-TGYTSQVSGRGKLTDFPDKG 96
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 69 LFLVEVTOLTESDGGVYACGAGMNTDRGKTQKVTLVNHSHEYEP 111
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 97 EFVVTVDLQNTDNGSYKCGGVN-GRGLDFGVNVLVSQKPEP 138
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
RESULT 6
PIGR_PIG STANDARD; PRT; 102 AA.
AC Q29244;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Polymeric-immunoglobulin receptor (Poly-Ig receptor) (PIGR)
DE (Fragment).
GN PIGR.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Small intestine;
RX MEDLINE=96327607; PubMed=8672129;
RA Winteroe A.K., Fredholm M., Davies W.;
RT "Evaluation and characterization of a porcine small intestine cDNA
RT library: analysis of 839 clones.";
RL Mamm. Genome 7:509-517(1996).
CC -!- FUNCTION: THIS RECEPTOR BINDS POLYMERIC IGA AND IGM AT THE
CC BASOLATERAL SURFACE OF EPITHELIAL CELLS. THE COMPLEX IS THEN
CC TRANSPORTED ACROSS THE CELL TO BE SECRETED AT THE APICAL SURFACE.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. ALSO SECRETED.
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL: F14851; CAA23294.1; -.
DR InterPro: IPR003600; Ig_Like.
DR SMART: SM00410; Ig_Like; 1.
KW Immunoglobulin domain; Repeat; Transmembrane; Glycoprotein.
FT NON_TER 1 1
FT NON_TER 102 102
SQ SEQUENCE 102 AA; 11205 MW; 82C915264B1508E8 CRC64;
Query Match 9.6%; Score 120.5; DB 1; Length 102;
Best Local Similarity 35.2%; Pred. No. 0.0013;
Matches 31; Conservative 19; Mismatches 31; Indels 7; Gaps 5;
QY 4 PEVKEGELGGSVTIKCPLPE--MHVRIYLCREMAGSGTCGTVVSTTNFKAEYKGRVTL 61
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 9 PEL-IVGDLRGSVTFDCALGQEMANVAKFLC-QLKNGKTCNVVINTLGKKAQDFEGRILL 66
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 62 KQYPRKNL-FLVEVTOLTESDGGVYACG 88
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 67 T--PRENSHFSVHITGLRKEDAGHYLCG 92
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
RESULT 7
CM35_HUMAN STANDARD; PRT; 224 AA.
ID CM35_HUMAN
AC Q08708;
DT 01-OCT-1996 (Rel. 34, Created)

```

DT	01-OCT-1996 (Rel. 34, Last sequence update)	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
DT	01-MAR-2002 (Rel. 41, Last annotation update)	OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
DE	CMRF35 antigen precursor	OX	NCBI_TaxID=9606;
GN	CMRF35.	RN	[1]
OS	Homo sapiens (Human).	RN	SEQUENCE FROM N.A. (CD22-BETA).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	RC	TISSUE=Tonsil;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	RX	MEDLINE=91086838; PubMed=1985119;
OX	NCBI_TaxID=9606;	RA	Wilson G.L., Fox C.H., Fauci A.S., Kehrl J.H.;
RN	[1]	RT	*cDNA cloning of the B cell membrane protein CD22: a mediator of B-B
RP	SEQUENCE FROM N.A.	RT	cell interactions.;
RX	MEDLINE=92249405; PubMed=1349532;	RL	J. Exp. Med. 173:137-146(1991).
RA	Jackson D.G., Hart D.N.J., Starling G., Bell J.I.;	RN	[2]
RT	*Molecular cloning of a novel member of the immunoglobulin gene	RX	SEQUENCE FROM N.A. (CD22-BETA).
RT	superfamily homologous to the polymeric immunoglobulin receptor.;	RX	MEDLINE=93267103; PubMed=8496602;
RL	Eur. J. Immunol. 22:1157-1163(1992).	RA	Wilson G.L., Najfeld V., Kozlow E., Menniger J., Ward D.,
CC	-1- SUBCELLULAR LOCATION: Type 1 membrane protein (Potential).	RA	Kehrl J.H.;
CC	-1- TISSUE SPECIFICITY: PRESENT ON THE SURFACE OF MONOCYTES,	RT	*Genomic structure and chromosomal mapping of the human CD22 gene.;
CC	NEUTROPHILS, A PROPORTION OF PERIPHERAL BLOOD T AND B LYMPHOCYTES	RL	J. Immunol. 150:5013-5024(1993).
CC	AND LYMPHOCYTIC CELL LINES.	RN	[3]
CC	-1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.	RX	SEQUENCE FROM N.A. (CD22-BETA).
CC	-1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.	RA	Lamerdin J.E., McCready P., Adamson A.W., Burkhart-Schultz K.,
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	RA	Garcia E., Kyle A., Ramirez M., Stilwagen S., Gaines J.,
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	RA	Dangnan L., Bruce R., Quan G., Montgomery M., Ow D.,
CC	the European Bioinformatics Institute. There are no restrictions on its	RA	Kobayashi A., Olsen A.O., Carrano A.V.;
CC	use by non-profit institutions as long as its content is in no way	RL	Submitted (AUG-1996) to the EMBL/GenBank/DBAJ databases.
CC	modified and this statement is not removed. Usage by and for commercial	RN	[4]
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/	RX	SEQUENCE FROM N.A. (CD22-ALPHA).
CC	or send an email to license@isb-sib.ch).	RX	MEDLINE=90231465; PubMed=1691828;
CC	-----	RA	Stamenkovic I., Seed B.;
CC	EMBL; X66171; CAA46948.1; -.	RT	*The B-cell antigen CD22 mediates monocyte and erythrocyte adhesion.;
DR	InterPro; IPR003599; Ig.	RL	Nature 345:74-77(1990).
DR	InterPro; IPR003006; Ig_MHC.	RN	[5]
DR	Pfam; PF00047; Ig; 1.	RX	VARIANTS GLU-152; GLY-664; CYS-669 AND ASP-745.
DR	SMART; SM00409; IG; 1.	RX	MEDLINE=99180618; PubMed=10079291;
KW	Antigen; Transmembrane; Glycoprotein; Signal.	RA	Hatta Y., Tsuchiya N., Matsushita M., Shiota M., Hagiwara K.,
FT	SIGNAL	RA	Tokunaga K.;
FT	CHAIN	RT	*Identification of the gene variations in human CD22.;
FT	DOMAIN	RL	Immunogenetics 49:280-286(1999).
FT	TRANSMEM	CC	-1- FUNCTION: MEDIATES B-CELL B-CELL INTERACTIONS. MAY BE INVOLVED IN
FT	DOMAIN	CC	THE LOCALIZATION OF B-CELLS IN LYMPHOID TISSUES. BINDS STALYLATED
FT	DOMAIN	CC	GLYCOPROTEINS: ONE OF WHICH IS CD45.
FT	DOMAIN	CC	-1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
FT	DISULFID	CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.
FT	DISULFID	CC	HERE; ARE PRODUCED BY ALTERNATIVE SPLICING.
FT	CARBOHYD	CC	-1- TISSUE SPECIFICITY: B-LYMPHOCYTES.
FT	CARBOHYD	CC	-1- PTM: PHOSPHORYLATED BOTH ON THREONINE/SERINE AND TYROSINE.
FT	CARBOHYD	CC	-1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
SQ	SEQUENCE 224 AA; 24830 MW; 60C88716D84600D2 CRC64;	CC	-1- SIMILARITY: CONTAINS 6 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
		CC	-1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
		CC	-1- DATABASE: NAME=PROW; NOTE=CD guide CD22 entry;
		CC	WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd22.htm"
		CC	-----
QY	8 VEGELGSGVTIKCPLEWH--VRVYLCREMAGSGTCGVTVSTNFKAEYGRVTKQYP 65	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
DB	31 VAGPVGSLVQCYRKEHRTLNKFCRP-PQILRCDKIVETKG-SAGKRGVRSIDSP 88	CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
QY	66 RKNLFLVEVQLTSDSGVYACG 88	CC	the European Bioinformatics Institute. There are no restrictions on its
DB	89 ANLSFTVTLENLTEDAGTYWCG 111	CC	use by non-profit institutions as long as its content is in no way
		CC	modified and this statement is not removed. Usage by and for commercial
		CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
		CC	or send an email to license@isb-sib.ch).
		CC	-----
RESULT 8		CC	EMBL; X59350; CAA42006.1; -.
CD22_HUMAN		DR	EMBL; X59350; CAA42006.1; -.
ID	CD22_HUMAN STANDARD; PRT: 847 AA.	DR	EMBL; X52785; CAA36988.1; -.
AC	P20273; Q01665; Q92872; O95699; O95701; O95702; O95703;	DR	EMBL; AB012997; BAA36566.1; -.
DT	01-FEB-1991 (Rel. 17, Created)	DR	EMBL; AB013003; BAA36572.1; -.
DT	15-JUL-1999 (Rel. 38, Last sequence update)	DR	EMBL; AB013004; BAA36573.1; -.
DT	01-MAR-2002 (Rel. 41, Last annotation update)	DR	EMBL; AB013006; BAA36575.1; -.
DE	B-cell receptor CD22 precursor (Leu-14) (B-lymphocyte cell adhesion	DR	PIR; A35648; A35648.
DE	molecule) (BL-CAM).	DR	PIR; JH0371; JH0371.
GN	CD22.	DR	MIM; 107266; -.
OS	Homo sapiens (Human).	DR	InterPro; IPR003006; Ig_MHC.
		DR	InterPro; IPR003598; Ig_C2.

Db 229 I-----STNQSHPSAVSTANT-PGSPNEALKQLRASKEDSCERILKNAMKRRHLADQDW 282
 QY 218 RQAL--DYGSQ 227
 Db 283 RQVLVVICGQD 294

RESULT 10
 NCRL_HUMAN
 ID NCRL_HUMAN STANDARD; PRT; 2440 AA.
 AC 075376; O9U018;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Nuclear receptor co-repressor 1 (N-CoR1) (N-CoR).
 GN NCRL OR KIAA1047.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal brain;
 RX MEDLINE=98393736; PubMed=9724795;
 RA Wang J., Hoshino T., Redner R.L., Kajigaya S., Liu J.M.;
 RT "ETO, fusion partner in t(8;21) acute myeloid leukemia, represses
 transcription by interaction with the human N-CoR/MSIN3/HDAC1
 complex.";
 RT Proc. Natl. Acad. Sci. U.S.A. 95:10860-10865(1998).
 RL [2]
 RN SEQUENCE OF 782-2440 FROM N.A.
 RP TISSUE=Brain;
 RX MEDLINE=99397452; PubMed=10470851;
 RA Kikuno R., Nagase T., Ishikawa K.-I., Hirose M., Miyajima N.,
 RA Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XIV.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 6:197-205(1999).
 RN [3]
 RP SEQUENCE OF 974-2440 FROM N.A.
 RX MEDLINE=99375328; PubMed=1044336;
 RA Nagaya T., Chen K.-S., Fujieda M.,
 RA Horwitz K.B., Lupski J.R., Seo H.;
 RT "Localization of the human nuclear receptor co-repressor (hn-CoR) gene
 RT between the CMT1A and the SMS critical regions of chromosome
 RT 17p11.2.";
 RL Genomics 59:339-341(1999).
 CC -|- FUNCTION: MEDIATES THE TRANSCRIPTIONAL REPRESSION ACTIVITY OF SOME
 CC NUCLEAR RECEPTORS BY PROMOTING CHROMATIN CONDENSATION, THUS
 CC PREVENTING ACCESS OF THE BASAL TRANSCRIPTION.
 CC -|- SUBUNIT: FORMS A LARGE COREPRESSOR COMPLEX THAT CONTAINS SIN3A/B
 CC AND HISTONE DEACETYLASES HDAC1 AND HDAC2. THIS COMPLEX ASSOCIATES
 CC WITH THE THYROID (TR) AND THE RETINOIC ACID RECEPTORS (RAR) IN THE
 CC ABSENCE OF LIGAND.
 CC -|- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -|- DOMAIN: THE N-TERMINAL REGION CONTAINS REPRESSION FUNCTIONS THAT
 CC ARE DIVIDED INTO THREE INDEPENDANT REPRESSION DOMAINS (RD1, RD2
 CC AND RD3). THE C-TERMINAL REGION CONTAINS THE NUCLEAR RECEPTOR-
 CC INTERACTING DOMAINS THAT ARE DIVIDED IN TWO SEPARATE INTERACTION
 CC DOMAINS (ID1 AND ID2).
 CC -|- DOMAIN: THE TWO INTERACTION DOMAINS (ID) CONTAIN A CONSERVED
 CC SEQUENCE REFERRED TO AS THE CORNR BOX. THIS MOTIF IS REQUIRED AND
 CC SUFFICIENT TO PERMIT BINDING TO UNLIGANDED TR AND RARS. SEQUENCES
 CC FLANKING THE CORNR BOX DETERMINE NUCLEAR HORMONE RECEPTOR
 CC SPECIFICITY.
 CC -|- SIMILARITY: CONTAINS 1 SANT-A DOMAIN.
 CC -|- SIMILARITY: CONTAINS 1 MYB-LIKE DOMAIN.
 CC -|- SIMILARITY: CONTAINS 2 CORNR BOX.
 CC -|- SIMILARITY: BELONGS TO THE N-COR NUCLEAR RECEPTOR COREPRESSORS
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF044209; AAC33550.1; -;
 DR EMBL; AF028970; BAA82999.1; -;
 DR EMBL; AB019524; BAA75814.1; -;
 DR MIM; 600849; -;
 DR InterPro; IPR001005; Myb_DNA_bind.
 DR Pfam; PF00249; myb_DNA-binding; 2.
 DR SMART; SM00395; SANT; 2.
 DR PROSITE; PS00300; MYB_3; 1.
 KW Nuclear protein; Transcription regulation; DNA-binding; Repressor;
 KW Coiled coil.
 FT DOMAIN 174 216 COILED COIL (POTENTIAL).
 FT DOMAIN 254 312 INTERACTION WITH SIN3A/B.
 FT DOMAIN 299 328 COILED COIL (POTENTIAL).
 FT DNA_BIND 437 482 SANT-A (POTENTIAL).
 FT DNA_BIND 620 670 MYB.
 FT DOMAIN 501 557 COILED COIL (POTENTIAL).
 FT DOMAIN 607 617 PRO-RICH.
 FT DOMAIN 988 1816 INTERACTION WITH ETO.
 FT DOMAIN 2055 2059 CORNR BOX OF ID1.
 FT DOMAIN 2263 2267 CORNR BOX OF ID2.
 FT DOMAIN 58 64 POLY-GLN.
 FT DOMAIN 593 603 POLY-ALA.
 FT DOMAIN 1032 1035 POLY-PRO.
 FT DOMAIN 1707 1712 POLY-ALA.
 FT DOMAIN 1952 1963 POLY-SER.
 FT CONFLICT 1014 1014 L -> V (IN REF. 2).
 FT CONFLICT 1508 1509 PP -> SS (IN REF. 2).
 FT CONFLICT 1561 1561 W -> R (IN REF. 2).
 FT CONFLICT 1567 1567 Q -> H (IN REF. 2).
 SQ SEQUENCE 2440 AA; 270263 MW; 60A4D7964D00EDAB CRC64;
 Query Match 7.48; Score 93; DB 1; Length 2440;
 Best Local Similarity 23.08; Pred. No. 12;
 Matches 50; Conservative 24; Mismatches 79; Indels 64; Gaps 10;
 QY 14 GSVTIKCP-----LPENHVIYLCREMGSGTC---GTVVS---TTNFIAEYKGR 58
 Db 912 GSILVSSPLKPNLPLQLQHRAAVIPPVMSVCTPCNIPGTPVSGYALYQRHAKAMHESA 971
 QY 59 VTLKQYPRKNLFLVEVTQLTESDSGVYACGAGNMTRDGTQKVTLVNHSEYFSEWEE-QP 117
 Db 972 LLEORORQEOIDECRSSTS-----PCGTSKSPNR-----EWEVLQP 1009
 QY 118 MP-----ETPKWFHLPYLFQMPAYASSSKFVTRVTPAQRGVPPVHHSPPTQIHRPR 172
 Db 1010 APhOLITNLPGVRLP-----TTRPTRPP-----PPLIPSSKTTVASEKPS 1050
 QY 173 VSRASSVAGDKPRTFLPSTTASKISALEGLLKPTQPS 209
 Db 1051 FIMGSGISQGTPTGYL--TSHNQASYTQETPKPSVGS 1085
 RESULT 11
 PLK_CHICK
 ID PLK_CHICK STANDARD; PRT; 355 AA.
 AC P07354;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Proteoglycan link protein precursor (Cartilage link protein) (LP).
 GN CRT11.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;

Tue Oct 29 09:47:51 2002

```
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J04374; AAA43039.1; -.
DR FIR; J00102; RKPPEM.
DR MEROPS; C21.001; -.
DR InterPro; IPR001788; RNA_dep_RNapol2.
DR InterPro; IPR000606; Viral_helicase1.
DR Pfam; PF00978; RNA_dep_RNapol2; 1.
DR Pfam; PF01443; Viral_helicase1; 1.
DR Transferase; RNA-directed RNA polymerase; Polyprotein; ATP-binding.
KW NP_BIND 965 972 ATP (BY SIMILARITY).
FT SEQUENCE 1839 AA; 204731 MW; FD8DC1F5115E7861 CRC64;
SQ
Query Match 7.2%; Score 90.5; DB 1; Length 1839;
Best Local Similarity 26.0%; Pred. No. 14;
Matches 46; Conservative 16; Mismatches 80; Indels 35; Gaps 7;
Qy 58 RYTLKQYPRKNL-----FLVEVTQLTESDGVYACGAGMNTDRGKTOKVTLNH-SEYEP 111
Db 493 RIELQRLPLMSLIPKIVLPILLSLLSSPTIYIHFFQAQTPQQLHDNYHLHPSRFEL 552
Qy 112 SWEQPMPTPKWPHLPYLFQMPAYASSKXEVTRVTPAQRCKYPPVHH----- 160
Db 553 SWTLQSYHVTQAQSLPQLLLPAPTQAQAS-----NPAPR---PPAFHAIPLPQFPTS 602
Qy 161 SSFTTQ-----ITRRPRVSRASSVAGDK-PRTFPLPSTTASKISALEGLLKQPTS 209
Db 603 SSPPLQEPPLSPHLIHPPLTREPSPLNGACDSALLPSTAAMTSAEHPTPLNPPTS 659
```

Search completed: October 28, 2002, 17:31:27
Job time : 10.6613 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2002, 17:24:41 ; Search time 3.33065 Seconds
(without alignments)
1371.778 Million cell updates/sec

Title: US-09-135-238B-2_COPY_273_390
Perfect score: 616
Sequence: 1 KRAVERKALSRARRLAVR.....HQPAAAMMEDSDDDYINPVA 118

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	98.5	16.0	116	1 STP2_RAT	P11101 rattus norv
2	87	14.1	2045	1 CBP1_CAEEL	P34545 caenorhabdi
3	85.5	13.9	2205	1 POLN_RUBVT	P13889 rubella vir
4	85.5	13.9	3866	1 HRX_MOUSE	P55200 mus musculu
5	81	13.1	549	1 FAST_HUMAN	Q14296 homo sapien
6	81	13.1	550	1 SRCE_HUMAN	Q14247 homo sapien
7	80	13.0	466	1 BIAR_MOUSE	P34971 mus musculu
8	80	13.0	466	1 BIAR_RAT	P18090 rattus norv
9	79.5	12.9	525	1 BTB2_HUMAN	Q9bx70 homo sapien
10	79	12.8	295	1 PLS3_HUMAN	Q9ncv6 homo sapien
11	78.5	12.7	425	1 POYN_DROME	P23758 drosophila
12	78	12.7	863	1 MCMA_HUMAN	P33991 homo sapien
13	77.5	12.6	415	1 SYNT_CANFA	O62732 canis famli
14	77.5	12.6	942	1 M3KE_MOUSE	Q9wu16 mus musculu
15	77	12.5	248	1 ICT3_HSV11	P36313 herpes simp
16	77	12.5	577	1 CST2_HUMAN	P33240 homo sapien
17	77	12.5	938	1 EBNA_EBY	P03203 epstein-bar
18	76.5	12.4	258	1 UL49_HSVBC	P30022 bovine herp
19	76.5	12.4	444	1 FXF2_HUMAN	Q12947 homo sapien
20	76.5	12.4	587	1 FOLC_MOUSE	P48760 mus musculu
21	76.5	12.4	862	1 MCMA_MOUSE	P49717 mus musculu
22	76	12.3	1185	1 DRPL_HUMAN	P54259 homo sapien
23	75.5	12.3	988	1 CLC1_HUMAN	P35523 homo sapien
24	75.5	12.3	1616	1 SLAP_BACCI	P35824 bacillus cl
25	75	12.2	285	1 HXA4_MOUSE	P06798 mus musculu
26	75	12.2	335	1 PLS1_RAT	P58195 rattus norv
27	75	12.2	495	1 SMA6_MOUSE	Q35182 mus musculu
28	75	12.2	543	1 VP61_NPVAC	Q03209 autographa
29	75	12.2	994	1 CLC1_MOUSE	Q64347 mus musculu
30	75	12.2	1271	1 BCR_HUMAN	P11274 homo sapien
31	74.5	12.1	294	1 YQ19_CAEEL	Q09507 caenorhabdi
32	74.5	12.1	308	1 CC40_CAEEL	P34804 caenorhabdi
33	74.5	12.1	366	1 IHA_RAT	P17490 rattus norv

RESULT 1

STP2_RAT

ID	STP2_RAT	STANDARD:	PRT:	116 AA.
AC	P11101:			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	01-JUL-1989 (Rel. 11, Last sequence update)			
DT	01-FEB-1996 (Rel. 33, Last annotation update)			
DE	Nuclear transition protein 2 (Tp-2).			
GN	TNP2.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89263797; PubMed=2726489;			
RA	Luerssen H., Maier W.M., Hoyer-Fender S., Engel W.;			
RT	"The nucleotide sequence of rat transition protein 2 (TP2) cDNA.";			
RL	Nucleic Acids Res. 17:3585-3585(1989).			
RN	[2]			
RP	SEQUENCE OF 91-116.			
RX	MEDLINE=87326384; PubMed=3307778;			
RA	Cole K.D., Kistler W.S.;			
RT	"Nuclear transition protein 2 (TP2) of mammalian spermatids has a			
RL	very basic carboxyl terminal domain.";			
RN	[3]			
RP	ZINC-BINDING.			
RX	MEDLINE=92028893; PubMed=1930189;			
RA	Baskaran R., Rao M.R.S.;			
RT	"Mammalian spermatid specific protein, TP2, is a zinc metalloprotein			
RL	with two finger motifs.";			
RN	[4]			
CC	BIOCHEM. BIOPHYS. RES. COMMUN. 179:1491-1499(1991).			
CC	FUNCTION: IN THE ELONGATING SPERMATIDS OF MAMMALS, THE CONVERSION			
CC	OF NUCLEOSOMAL CHROMATIN TO THE COMPACT, NONNUCLEOSOMAL FORM FOUND			
CC	IN THE SPERM NUCLEUS IS ASSOCIATED WITH THE APPEARANCE OF A SMALL			
CC	SET OF BASIC CHROMOSOMAL TRANSITION PROTEINS.			
CC	!- SUBCELLULAR LOCATION: Nuclear.			
CC	!- SIMILARITY: STRONG, TO OTHER MAMMALIAN SPERMATID NUCLEAR			
CC	TRANSITION PROTEINS 2.			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; X14776; CAA32882.1; .			
DR	PIR; A26834; A26834.			
DR	PIR; S04094; S04094.			
DR	InterPro; IPR000678; TP2.			
DR	Pfam; PF01254; TP2; 1.			
DR	PROSITE; PS00970; TP2_1; 1.			
DR	PROSITE; PS00971; TP2_2; 1.			

34	74	12.0	583	1	ESR1_OREAU	P50240 oreochromis
35	74	12.0	669	1	COTE_HUMAN	P81408 homo sapien
36	74	12.0	763	1	SAT1_HUMAN	Q01826 homo sapien
37	74	12.0	764	1	SAT1_MOUSE	Q06611 mus musculu
38	74	12.0	823	1	YRR2_CAEEL	Q09345 caenorhabdi
39	74	12.0	1496	1	CA25_HUMAN	P05997 homo sapien
40	74	12.0	3969	1	HRX_HUMAN	Q03164 homo sapien
41	73.5	11.9	199	1	AA27_MOUSE	P56873 mus musculu
42	73.5	11.9	292	1	YO14_CAEEL	Q09505 caenorhabdi
43	73.5	11.9	467	1	VE2_HPV24	P50770 human papil
44	73.5	11.9	522	1	PLIN_HUMAN	O60240 homo sapien
45	73.5	11.9	590	1	FTZ1_RHINE	P30327 rhizobium m

ALIGNMENTS


```

DR Pfam; PF00978; RNA_dep_RNApol2; 1.
DR SMART; SM00506; Alpp; 1.
KW Polypeptide; Nonstructural protein.
FT CHAIN 1 ? NONSTRUCTURAL PROTEIN NSP1.
FT CHAIN 2 ? NONSTRUCTURAL PROTEIN NSP2.
FT CHAIN 3 ? NONSTRUCTURAL PROTEIN NSP3.
FT CHAIN 4 ? NONSTRUCTURAL PROTEIN NSP4.
FT CHAIN 5 ? NONSTRUCTURAL PROTEIN NSP5.
SQ SEQUENCE 2205 AA; 240220 MW; 15A95F5E34C0B03 CRC64;

Query Match 13.9%; Score 85.5; DB 1; Length 2205;
Best Local Similarity 31.4%; Pred. No. 14;
Matches 38; Conservative 9; Mismatches 43; Indels 31; Gaps 7;

QY 9 ALSRRRLAVRM--RALESSQPRGSPRRSONNIYACPRRA-----RGADAAGT---G 59
Db AFANAVTAAVRAGPRQSAASAPPGDPPP-----PRRRSRHSDARGTTPPA 748

QY 60 EAPVPGCAPLPAPLQVSE--SPWLHAPS-----LKTCEYVSLYHQPAAMMEDSDS 110
Db FARDDPPPPAPPPAPPRAGDPVPPIPAGPADRADAELEVACPSG---PPTSTRADPDS 805

QY 111 D 111
Db 806 D 806

RESULT 4
ID HRX_MOUSE STANDARD; PRT; 3866 AA.
AC P55200;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Zinc finger protein HRX (ALL-1) (Fragment).
GN MLL OR HRX OR ALL1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC STRAIN=C57BL/6J, AND B6/CBA; TISSUE=Spleen, and Lung;
RA MEDLINE=93317679; PubMed=8327517;
RA Ma Q., Alder H., Nelson K.K., Chatterjee D., Gu Y., Nakamura T.,
RA Canani E., Croce C.M., Siracusa L.D., Buchberg A.M.;
RT "Analysis of the murine All-1 gene reveals conserved domains with
RT human All-1 and identifies a motif shared with DNA
RT methyltransferases.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:6350-6354(1993).
CC -!- FUNCTION: POSSIBLY ACTS AS A TRANSCRIPTIONAL REGULATORY FACTOR.
CC MAY REGULATE GENES INVOLVED IN SKELETAL FORMATION DURING
CC EMBRYONIC DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -!- SIMILARITY: BELONGS TO THE TRANSCRIPTION FACTOR TRITHORAX FAMILY.
CC -!- SIMILARITY: CONTAINS 1 BROMODOMAIN.
CC -!- SIMILARITY: CONTAINS 1 SET DOMAIN.
CC -!- SIMILARITY: CONTAINS 3 PHD-TYPE ZINC FINGERS.
CC -!- SIMILARITY: CONTAINS 1 CXXC-TYPE ZINC FINGER.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L17069; AAA62593.1; -.
CC MGD; MGI:96995; M11.
DR InterPro; IPR001487; Bromodomain.
DR InterPro; IPR003889; FYrich_C.

```

```

DR InterPro; IPR003888; FYrich_N.
DR InterPro; IPR001965; PHD.
DR InterPro; IPR003636; PostSET.
DR InterPro; IPR000093; RecR.
DR InterPro; IPR001214; SET.
DR InterPro; IPR002857; Znf-CXXC.
DR Pfam; PF00628; PHD; 3.
DR Pfam; PF00856; SET; 1.
DR Pfam; PF02008; zf-CXXC; 1.
DR SMART; SM00297; BROMO; 1.
DR SMART; SM00542; FYRC; 1.
DR SMART; SM00541; FYRN; 1.
DR SMART; SM00249; PHD; 4.
DR SMART; SM00508; PostSET; 1.
DR SMART; SM00317; SET; 1.
DR PROSITE; PS00014; BROMODOMAIN_2; 1.
DR PROSITE; PS0280; SET; 1.
KW DNA-binding; Nuclear protein; Zinc-finger; Metal-binding;
KW Transcription regulation; Alternative splicing; Polymorphism.
FT NON_TER 1 1
FT DNA_BIND 67 78 A.T HOOK (BY SIMILARITY).
FT DNA_BIND 115 125 A.T HOOK (BY SIMILARITY).
FT DNA_BIND 199 207 A.T HOOK (BY SIMILARITY).
FT ZN_FING 1044 1091 CXXC-TYPE.
FT ZN_FING 1330 1381 PHD-TYPE 1.
FT ZN_FING 1383 1432 PHD-TYPE 2.
FT ZN_FING 1465 1529 PHD-TYPE 3.
FT DOMAIN 1605 1650 BROMODOMAIN (DIVERGENT).
FT DOMAIN 3737 3846 SET.
FT DOMAIN 35 41 POLY-GLY.
FT DOMAIN 459 469 PRO-RICH.
FT DOMAIN 1231 1238 POLY-PRO.
FT DOMAIN 3533 3536 POLY-GLU.
FT DOMAIN 3693 3697 POLY-GLU.
FT VARSPIC 1503 1505 MISSING (IN ISOFORM 2).
FT VARIANT 1497 1497 K -> T.
SQ SEQUENCE 3866 AA; 420976 MW; ADFC55E14E806FID CRC64;

Query Match 13.9%; Score 85.5; DB 1; Length 3866;
Best Local Similarity 28.2%; Pred. No. 24;
Matches 31; Conservative 21; Mismatches 37; Indels 21; Gaps 6;

QY 1 KLAVERRRKALSRRLAVRMRA-LESSQ---PRGSPRRSONNIYACPRRARGADA 55
Db 1111 KKEKESKTKTEKESKESVSKPLPAQKAAPPPREEAPKSS---SEPPRRKPVEEK 1167

QY 56 AGTGEAPVPGCAPLPAPLQVSESPWLHAPSLSKTCEYVSLYHQPAAM 105
Db 1168 SEEGGAPAPAPA---PEPKQVS-----APASRKSKQVS---QPAAVV 1204

RESULT 5
FAST_HUMAN STANDARD; PRT; 549 AA.
ID FAST_HUMAN
AC Q14296;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Fas-activated serine/threonine kinase (EC 2.7.1.-) (FAST kinase).
GN FASTK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Peripheral blood lymphocytes;
RX MEDLINE=95378805; PubMed=7544399;
RA Tian Q., Taupin J.-L., Elledge S., Robertson M., Anderson P.;
RT "Fas-activated serine/threonine kinase (FAST) phosphorylates TIA-1
RT during Fas-mediated apoptosis.";
RL J. Exp. Med. 182:865-874(1995).
CC -!- FUNCTION: FAS-MEDIATED APOPTOSIS IS CHARACTERIZED BY FASTK-

```


QY 78 SESPWLHAPSLKTCSEYV-----SLYHQPAAMME 106
 DB 62 SAAPFLPLGVSPGSLFVQIDQILHQAERVE 95

RESULT 11
 POXN_DROME
 AC P23758; STANDARD; PRT; 425 AA.
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 GN POX-N.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP MEDLINE=92208941; PubMed=1348214;
 RA Dambly-Chaudiere C., Jamet E., Burri M., Boop D., Basler K.,
 RA Hafen E., Dumont N., Spielmann P., Ghysen A., Noll M.;
 RT "The paired box gene pox-neuro: a determinant of poly-innervated
 sense organs in Drosophila.";
 RL Cell 69:159-172(1992).
 RN [2]
 RP MEDLINE=90059940; PubMed=2573516;
 RA Bopp D., Jamet E., Baumgartner S., Burri M., Noll M.;
 RT "Isolation of two tissue-specific Drosophila paired box genes, Pox
 meso and Pox-neuro.";
 RL EMBO J. 8:3447-3457(1989).
 RN [3]
 RP FUNCTION.
 RX MEDLINE=94121914; PubMed=8292359;
 RA Nottebohm E., Usui A., Therianos S., Kimura K., Dambly-Chaudiere C.,
 RA Ghysen A.;
 RT "The gene poxn controls different steps of the formation of
 chemosensory organs in Drosophila.";
 RL Neuron 12:25-34(1994).
 RN [4]
 RP FUNCTION.
 RX MEDLINE=97326936; PubMed=9183748;
 RA Awasaki T., Kimura K.;
 RT "pox-neuro is required for development of chemosensory bristles in
 Drosophila.";
 RL J. Neurobiol. 32:707-721(1997).
 CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT SPECIFIES POLY-INNERVATED
 CC ORGANS (CHEMOSENSORY BRISTLE). ALSO CONTROLS THE NUMBER OF
 CC NEURONS.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- TISSUE SPECIFICITY: CENTRAL AND PERIPHERAL NERVOUS SYSTEM.
 CC -!- SIMILARITY: CONTAINS A PAIRED BOX DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M86927; AAA28832.1; -;
 CC EMBL; X58917; CAA41721.1; -;
 CC PIR; S06951; S06951.
 CC HSP; P26367; 6PAX.
 CC FlyBase; FBgn0003130; Pox-n.
 CC InterPro; IPR001523; Paired_box.
 CC Pfam; PF00292; PAX; 1.
 CC PRINTS; PR00027; PAIREDBOX.

DR SMART; SM00351; PAX; 1.
 DR PROSITE; PS00034; PAIRED_BOX; 1.
 KW Paired box; Developmental protein; Nuclear protein;
 KW Transcription regulation; Neurogenesis; Differentiation.
 FT DOMAIN 5 130 PAIRED_BOX.
 FT DOMAIN 146 157 POLY-ALA.
 FT DOMAIN 173 176 POLY-PRO.
 FT DOMAIN 281 287 POLY-ALA.
 SQ SEQUENCE 425 AA; 44387 MW; EC76C6494A06DC15 CRC64;
 Query Match 12.7%; Score 78.5; DB 1; Length 425;
 Best Local Similarity 32.4%; Pred. NO. 9.7; Indels 2;
 Matches 23; Conservative 9; Mismatches 30; Gaps 2;
 QY 38 SNNIYSACPRRARGADAAGTGEAPVPGFAPLPAPLQVSESPWLHAPSLKTCSEYVSL 97
 DB 142 SQNAAAAAHHAGSG--PSNGYGGQAPPVTV-----APPTPAATPSIAR 192
 QY 98 YHQPAAMMEDS 108
 DB 193 YAKPPALMMNS 203

RESULT 12
 MCM4_HUMAN
 ID MCM4_HUMAN STANDARD; PRT; 863 AA.
 AC P33991; Q99658;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DNA replication licensing factor MCM4 (CDC21 homolog) (P1-CDC21).
 GN MCM4 OR CDC21
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hu B.;
 RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1-712 FROM N.A.
 RX MEDLINE=98126438; PubMed=9465298;
 RA Connelly M.A., Zhang H., Kieleczawa J., Anderson C.W.;
 RT "The promoters for human DNA-PKcs (PRKDC) and MCM4: divergently
 RT transcribed genes located at chromosome 8 band q11.";
 RL Genomics 47:71-83(1998).
 RN [3]
 RP SEQUENCE OF 1-23 FROM N.A.
 RX MEDLINE=97430835; PubMed=9284934;
 RA Ladenburger E.M., Fackelmayer F.O., Hamelster H., Knippers R.;
 RT "MCM4 and PRKDC, human genes encoding proteins MCM4 and DNA-PKcs, are
 RT close neighbours located on chromosome 8q12->q13.";
 RL Cytogenet. Cell Genet. 77:268-270(1997).
 RN [4]
 RP SEQUENCE OF 440-863 FROM N.A.
 RC TISSUE=Cervix;
 RX MEDLINE=94089373; PubMed=8265339;
 RA Hu B., Burkhardt R., Schulte D., Musahl C., Knippers R.;
 RT "The P1 family: a new class of nuclear mammalian proteins related to
 RT the yeast Mcm replication proteins.";
 RL Nucleic Acids Res. 21:5289-5293(1993).
 CC -!- FUNCTION: INVOLVED IN THE CONTROL OF DNA REPLICATION.
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE MCM FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC

```

Db 1171 SEEGAPAPAPA-----PEPKQVS-----APASRKSKQVS---QPAVV 1207

RESULT 8
S38480
nonstructural protein - rubella virus
C/Species: rubella virus
C/Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 2
C/Accession: S38480
R:Gillam, S.
submitted to the EMBL Data Library, March 1993
A:Description: Nucleotide sequence of the nonstructural protein g
A:Reference number: S38480
A:Accession: S38480
A>Status: preliminary
A:Molecule type: genomic RNA
A:Residues: 1-2115 <GIL>
A:Cross-references: EMBL:X72393; NID:g410507; PIDN:CAA51087.1; P
C:Superfamily: rubella virus nonstructural polyprotein

Query Match 13.4%; Score 82.5; DB 2; Length 2115
Best Local Similarity 30.4%; Pred. No. 41;
Matches 38; Conservative 14; Mismatches 48; Indels 2

QY 2 RAVERRKALSRARRLAVRM--RALESSQRPGRSPRPSQNNIYSACPRRA-----PRARRSQ
Db 693 RDLPRGTAFTANAVTAARVAGPQLAATSPPGDPPP-----PPRRRRSQ
QY 56 AGT-GEAPVGP--GAPLPAPLQVSE--SPWLHAPSLKTSCEYVSLYHQPA---
Db 743 RGTPTTAPVVRDPPRPQPSPAPPVGVDPVPTTAAEPADRARHAELEVYVEPSGPP
QY 107 DSDSD 111
Db 803 DPDS D 807

RESULT 9
B75310
conserved hypothetical protein - Deinococcus radiodurans (strain H
C/Species: Deinococcus radiodurans
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 1
C/Accession: B75310
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococo
A:Reference number: A75250; MUID:20036896
A:Accession: B75310
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-528 <WHI>
A:Cross-references: GB:AE002048; GB:AE000513; NID:g6459929; PIDN:
A:Experimental source: Strain R1
C:Genetics:
A:Gene: DR2133
A:Map position: 1

Query Match 13.2%; Score 81.5; DB 2; Length 528;
Best Local Similarity 31.6%; Pred. No. 13;
Matches 30; Conservative 9; Mismatches 47; Indels 9

QY 2 RAVERRKALSRARRLAVRMRALESSQRPGRSPRPSQNNIYSACPRRARGADAAC
Db 437 RTSPRARASRASRPITIPAAANSASAAAPPNSPTRKTWNSTPGWCPRSAASTPSS
QY 62 PVP--GPGAPLPAPLQVSESPW--LHAPSLKTS C 92
Db 497 PPRPVGPG-PEPTA----RRRGWGRPHRAPKWKPN C 526

RESULT 10

```

I37386

fas-activated serine/threonine kinase - human

C:Species: Homo sapiens (man)

C>Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000

C:Accession: I37386

R:Rian, O.; Taupin, J.; Ellledge, S.; Robertson, M.; Anderson, P.

J. Exp. Med. 182, 865-874, 1995

A:Title: Fas-activated serine/threonine kinase (FAST) phosphorylates TIA-1 during Fas-me

A:Reference number: I37386; MUID:95378805

A:Accession: I37386

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-549 <RES>

A:Cross-references: EMBL:X86779; NID:g1006658; PIDN:CAA60448.1; PID:g1006659

C:Genetics:

A:Gene: fast

C:Superfamily: human fas-activated serine/threonine kinase

Query Match 13.1%; Score 81; DB 2; Length 549;
 Best Local Similarity 29.1%; Pred. No. 15;
 Matches 23; Conservative 5; Mismatches 23; Indels 28; Gaps 2;

QY 28 ORRGSPRRSQNNIYSACPRARGADAAGTGEAPVPGCAPL-----70

DB 2 RRRGPGPRA-----PRTEGATCAGGGSWSPSNMLRVLLSAQTSPARLSGLL 53

QY 71 ---PPAPLQVSESPWLHAP 86

DB 54 LIPPVQCCLGPGSKWGRP 72

RESULT 11

A48063

mammary tumor/squamous cell carcinoma-associated protein EMS1 - human

C:Species: Homo sapiens (man)

C>Date: 24-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 29-Aug-1997

C:Accession: A48063

R:Schuring, E.; Verhoeven, E.; Litvinov, S.; Michalides, R.J.

Mol. Cell. Biol. 13, 2891-2898, 1993

A:Title: The product of the EMS1 gene, amplified and overexpressed in human carcinomas,

A:Reference number: A48063; MUID:93233650

A:Accession: A48063

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-550 <SCH>

A:Note: sequence extracted from NCBI backbone (NCBIP:129872)

C:Genetics:

A:Gene: GDB:EMS1

A:Cross-references: GDB:202962; OMIM:164765

A:Map position: 11q13-11q13

C:Superfamily: SH3 homology

F:499-546/Domain: SH3 homology <SH3>

Query Match 13.1%; Score 81; DB 2; Length 550;
 Best Local Similarity 27.0%; Pred. No. 15;
 Matches 37; Conservative 16; Mismatches 56; Indels 28; Gaps 6;

QY 1 KRAVERKALSRARRLAVMRALSSOR-----PRGSPRRSQNNIYSACPR 48

DB 361 KEQEDRRKAEARQMAKERQEQEARRKLEQARAKTQTFPVSPAPQTEERLPSSFV 420

QY 49 RARGADAG-TGEAPVPGCAPLPPAPLQVSESPWLHAPSLKTSCEYV--SLY----- 98

DB 421 Y----EADASFKAELSYRGVSGTPEPEYISMEADYREASSQQLAYATEAVYESAEPG 477

QY 99 HQPAAAMDESDSDYIN 115

DB 478 HYP-----EDSYDEYEN 491

RESULT 12

C75318

hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000

C:Accession: C75318

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.

, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896

A:Accession: C75318

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-553 <WH>

A:Cross-references: GB:AE002044; GB:AE000513; NID:g6459872; PIDN:AAF11639.1; PID:g645

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR2090

A:Map position: 1

Query Match 13.1%; Score 81; DB 2; Length 553;
 Best Local Similarity 25.0%; Pred. No. 15;
 Matches 27; Conservative 16; Mismatches 37; Indels 28; Gaps 3;

QY 3 AVERKALSR---ARRLAVMRALSSORPRGSPRRSQNNIYSACPRRARGA----- 53

DB 176 AAPRREAPAPQVQERSTATQTQVQAAPKAPAPQPDSPSL----PRTLOALASDRL 231

QY 54 -----DAAGTGEAPVPGCAPLPPAPLQVSESPWLHAP 86

DB 232 PDLPLVELLRLWEQQAQEEEPAPAPAPVPVAAALSPAPSPAAP 279

RESULT 13

T12469

hypothetical protein DKFZp564C1940.1 - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 23-Jul-1999

C:Accession: T12469

R:Poustka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, June 1999

A:Reference number: 217525

A:Accession: T12469

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-296 <POU>

A:Cross-references: EMBL:AL080164

A:Experimental source: fetal brain; clone DKFZp564C1940

C:Genetics:

A:Note: DKFZp564C1940.1

Query Match 13.1%; Score 80.5; DB 2; Length 296;
 Best Local Similarity 28.3%; Pred. No. 9;
 Matches 43; Conservative 10; Mismatches 58; Indels 41; Gaps 6;

QY 6 RRKALSRARRLAVMR---ALESSORPRGSPRRSQNNIYSACPRRARGA 53

DB 128 RRRQRGLMRRLVHRLRWGLLPRTNTPARASEARSOVTPSAAPLEALDGGTGPAREGGA 187

QY 54 DAAGTGEAPVCP-GAPLP-----PAPLQVSESPWLHAPSLKTSCEYVSLYHQP----- 101

DB 188 VGSQDGEQAPLP.LKAPLSASTSPAPTTVPAPG-PLPSLPLEPSLLSGVQVQLGRLL 246

QY 102 -----AAMDESDSDYINVP 117

DB 247 PSLGPGPPTRRSPPGPHTAVLALEDDVLLVP 278

RESULT 14

T46482

hypothetical protein DKFZp434A045.1 - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000

C:Accession: T46482

R;Duesterhoeft, A.; Lauber, J.; Mewes, H.W.; Weil, B.; Wiemann, S.
 submitted to the Protein Sequence Database, January 2000
 A:Reference number: Z23036
 A:Accession: T46482
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-381 <AAA>
 A:Cross-references: EMBL:AL137629
 A:Experimental source: adult testis; clone DKFZp434A045
 C:Genetics:
 A:Note: DKFZp434A045.1

Query Match 13.1%; Score 80.5; DB 2; Length 381;
 Best Local Similarity 31.4%; Pred. No. 11;
 Matches 32; Conservative 12; Mismatches 27; Indels 31; Gaps 8;
 Qy 12 RRARLAVRRALESSQPRGSPRRPSONNIYSACPRRARGADAGTGEAPVPGP--GAP 69
 Db 227 RKERSTAV-MRS-QPARLPQASPRP-----YSSVP-----AGS-EKPPKGGSSYNPP 269
 Qy 70 LPPAPLQVSESPWLHAPSLKTCSEYVSLYHQPAAMMEDSDSD 111
 Db 270 LPPLKISTNSG---SPGFE-----YHQPCDKFEASKND 299

RESULT 15

T09024
 proline-rich protein T27E11.90 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 22-Oct-1999
 C:Accession: T09024
 R;Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancroft
 submitted to the Protein Sequence Database, June 1999
 A:Reference number: Z16533
 A:Accession: T09024
 A:Molecule type: DNA
 A:Residues: 1-577 <BEV>
 A:Cross-references: EMBL:AL078579; GSPDB:GN00062; ATSP:T27E11.90
 A:Experimental source: cultivar Columbia; BAC clone T27E11
 C:Genetics:
 A:Gene: ATSP:T27E11.90
 A:Map position: 4
 A:Introns: 26/1; 117/1; 338/1; 411/3; 430/2; 498/2

Query Match 13.1%; Score 80.5; DB 2; Length 577;
 Best Local Similarity 28.8%; Pred. No. 17;
 Matches 23; Conservative 10; Mismatches 32; Indels 15; Gaps 3;
 Qy 30 PRGSPRRPSONNIYSACPRRARGADAGTGEAPVPGGAPLP-----PAPLQVSE 79
 Db 234 FSSGTPGDSPLSPGPFSPSPPTGPDSPPLSPGDSPLSPGDPPLSPGPHLYEK 293
 Qy 80 SPWL-HAPSLKTCSEYVSLY 98
 Db 294 NRWLIHFPSIK-----YLSVF 309

Search completed: October 28, 2002, 17:34:06
 Job time : 10.1855 secs

GenCore version 5.1.3

Copyright (c) 1993 - 2002 CompuGen Ltd.
Source version 3.1.1.3

OM protein - protein search, using sw model

Run on: October 28, 2002, 17:24:40 ; Search time 40.8871 Seconds
(without alignments)
1059.473 Million cell updates/sec

Title: US-09-135-238B-2

Perfect score:

Sequence: 1 MDRWLWPLYFLPVSGALRIL.....HQPAAMMEDSDSDDYINVPA 390

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 100%

Listing first 45 summaries

Database :

Database : A_Geneseq_032802.*

1:	/SIDS1/qcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2:	/SIDS1/qcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3:	/SIDS1/qcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4:	/SIDS1/qcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5:	/SIDS1/qcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6:	/SIDS1/qcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7:	/SIDS1/qcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8:	/SIDS1/qcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9:	/SIDS1/qcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10:	/SIDS1/qcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11:	/SIDS1/qcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12:	/SIDS1/qcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13:	/SIDS1/qcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14:	/SIDS1/qcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15:	/SIDS1/qcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16:	/SIDS1/qcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17:	/SIDS1/qcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18:	/SIDS1/qcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19:	/SIDS1/qcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20:	/SIDS1/qcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21:	/SIDS1/qcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22:	/SIDS1/qcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	2047	99.6		390	20	AA12225	Human Toso protein
2	2047	99.6		390	20	AA17496	Human Toso protein
3	2047	99.6		390	20	AA105001	Human PIGRL-1 prot
4	1157	56.3		422	22	AAE05349	Mouse Toso protein
5	550	26.8		107	20	AA105002	Human PIGRL-1 prot
6	410	20.0		97	21	AA165401	Human 5' EST relat
7	198	9.6		771	17	AA103180	Mouse poly-immunog
8	195	9.5		771	22	AA165698	Mouse polymeric im
9	185	9.0		769	17	AA103181	Rat poly-immunoglob
10	185	9.0		769	22	AA165697	Rat polymeric immu
11	183.5	8.9		532	19	AA150033	Human immunity rel

12	182	8.9	607	20	AAY34099	Partial amino acid
13	182	8.9	607	20	AAW95601	Human secretory Im
14	182	8.9	746	17	AAW03178	Human poly-Immuno
15	182	8.9	757	17	AAW03179	Bovine poly-Immuno
16	182	8.9	757	22	AAG65696	Bovine polymeric i
17	182	8.9	764	22	AAG65695	Human polymeric im
18	182	8.9	764	22	AAG65711	Human polymeric im
19	179	8.7	733	22	AAG65699	Possum polymeric i
20	166.5	8.1	584	12	AAR14670	Truncated poly Ig
21	166.5	8.1	773	17	AAW03177	Rabbit poly-Immuno
22	166.5	8.1	773	22	AAG65700	Rabbit polymeric i
23	161	7.8	307	19	AAR38334	Rat kidney injury
24	156	7.6	584	12	AAR14671	Truncated poly Ig-
25	143.5	7.0	205	20	AAW05069	Human PI3K-2 prote
26	142.5	6.9	211	21	AAB15543	Human immune syste
27	131	6.4	746	22	AAU18104	Novel human secret
28	129.5	6.3	332	20	AAU14171	Human PRO846 prote
29	129.5	6.3	332	21	AAB44273	Human PRO846 (UNQ4
30	129.5	6.3	332	21	AAU19580	Human PRO846 used
31	129.5	6.3	332	22	AAU12364	Human PRO846 polyp
32	129.5	6.3	332	22	AAB65300	Human PRO846 prote
33	127.5	6.2	203	22	AAU14805	Novel bone marrow
34	127.5	6.2	313	22	AAB02639	Human DNAX surface
35	127.5	6.2	332	21	AAB27653	Human protein PRO8
36	122	5.9	298	20	AAU24023	Human CMRF-35-H9 r
37	121	5.9	334	19	AAW38336	Human kidney injur
38	120.5	5.9	301	20	AAU24022	Human CMRF-35-H9 r
39	118.5	5.8	528	15	AAU52952	Human anti-IgE Mab
40	117.5	5.7	476	18	AAW01818	PrimaTised anti-hu
41	117.5	5.7	476	19	AAB63761	Macaque primatized
42	117	5.7	453	14	AAR33311	Humanised MeaEl1 Ve
43	117	5.7	453	21	AAU05109	Heavy chain amino
44	117	5.7	1106	22	AAE06644	Human Cb1tuso inte
45	117	5.7	1106	22	AAB70611	Human GR1-1 amino

ALIGNMENTS

RESULT 1	
AA42225	
ID	AA42225 standard; Protein; 390 AA.
XX	
XX	
AC	AA42225;
XX	
XX	
DT	17-DEC-1999 (first entry)
XX	
DE	Human Toso protein sequence.
XX	
XX	
KW	Human; Toso protein; target; drug screening; diagnosis; apoptosis; apoptosis related disease.
KW	

XX The present invention describes a method of Screening for a bioactive
CC agent capable of binding a Toso protein. Also described a methods for:
CC (1) screening a bioactive agent capable of modulating activity of a
CC Toso cell-surface receptor, comprising adding a candidate bioactive
CC agent to a cell comprising a recombinant Toso nucleic acid, and
CC exposing the cells to an apoptotic agent that will induce apoptosis;
CC (2) modulating apoptosis comprising administering an exogenous
CC compound that binds Toso, to a cell; (3) identifying a cell containing
CC a mutant Toso gene, comprising determining it's sequence; (4)
CC identifying the Toso genotype, comprising determining the sequence of
CC at least one Toso gene; and (5) diagnosing an apoptosis related
CC condition, comprising measuring activity of Toso in a tissue, and
CC comparing to the activity from non-affected individual's tissue, where
CC a reduced activity of the patient indicates risk of an apoptosis related
CC condition. The methods are useful for identifying agents capable of
CC diagnosing and treating apoptosis related disease, their use for
CC modulating apoptosis, and methods for diagnosing the disease state.
CC The present sequence represents the human Toso protein for use in
CC methods from the invention.

XX Sequence 390 AA;

Query Match 99.6%; Score 2047; DB 20; Length 390;
Best Local Similarity 99.7%; Pred. No. 1.8e-166;
Matches 389; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDRWLWPLYFLPVSGALRLPEVKVEGELGSGVTIKCPLPEMHVRIYLCREMGAGSGTCT 60
|||
DB 1 MDFWLWPLYFLPVSGALRLPEVKVEGELGSGVTIKCPLPEMHVRIYLCREMGAGSGTCT 60

QY 61 VYSTNFIKAEYKGRVTLKQYPRKNLFLVEVTOLTESDSGVYACGAGMNTDRGKTOKVTL 120
|||
DB 61 VYSTNFIKAEYKGRVTLKQYPRKNLFLVEVTOLTESDSGVYACGAGMNTDRGKTOKVTL 120

QY 121 NVHSEYEPSWEQPMPTPKWFHLPYLFQMPAYASSSKFVTTPAQRGKVPVHHSSP 180
|||
DB 121 NVHSEYEPSWEQPMPTPKWFHLPYLFQMPAYASSSKFVTTPAQRGKVPVHHSSP 180

QY 181 TQOITHRPRVSRASSVAGDKPRTFLPSTTASKISALEGLLKQPTPSYNNHTRLHRQALD 240
|||
DB 181 TQOITHRPRVSRASSVAGDKPRTFLPSTTASKISALEGLLKQPTPSYNNHTRLHRQALD 240

QY 241 YGSGSGREGQGFHILPTILGLFLALLGLVVKRAVERRKALSRRLAVRRALESSQ 300
|||
DB 241 YGSGSGREGQGFHILPTILGLFLALLGLVVKRAVERRKALSRRLAVRRALESSQ 300

QY 301 RPRGSPRPRSONNIYSACPRRARGADAAGTGEAPVPGAPLPAPLQVSESPWLHAPSL 360
|||
DB 301 RPRGSPRPRSONNIYSACPRRARGADAAGTGEAPVPGAPLPAPLQVSESPWLHAPSL 360

QY 361 KTCEYVSLYHQPAAMMEDSDSDYINVPA 390
|||
DB 361 KTCEYVSLYHQPAAMMEDSDSDYINVPA 390

RESULT 2

AAV17496
ID AAV17496 standard; Protein; 390 AA.
XX AAV17496;
AC AAV17496;
XX 03-AUG-1999 (first entry)
XX Human Toso protein.
XX Toso protein; tumour necrosis factor mediated apoptosis inhibition;
KW TNF mediated apoptosis; T cell overactivity; autoimmune disease;
KW Sjogrens connective tissue disorder; transplant rejection; cancer.
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers

FT Peptide 1..17
FT /label= signal
FT Protein 18..390
FT /label= Toso
FT Region 254..272
FT /label= transmembrane_region

PN W09925832-A1.

XX 27-MAY-1999.

XX 16-NOV-1998; 98WO-US24391.

XX 17-AUG-1998; 98US-0135238.

PR 17-NOV-1997; 97US-0066063.

XX (STPD) UNIV LELAND STANFORD JUNIOR.

XX Hitoshi Y, Nolan GP;

PI WPI; 1999-338007/28.

DR N-PSDB; AAX76123.

XX DNA encoding Toso, a protein having inhibitory effects on TNF

PT mediated apoptosis

XX Claim 20; Fig 2a; 70pp; English.

XX The present sequence is a Toso protein (I). (I) has anti-apoptotic
CC and cytostatic activity. Toso (named after a Japanese liquor that is
CC drunk on New Year's Day to celebrate long life and eternal youth) most
CC likely acts by induction of cFLIP expression which inhibits caspase-8
CC processing. Recombinant (I) can be used to modulate apoptosis in a cell
CC or to treat an apoptosis related condition in a mammal. Apoptosis
CC related conditions can also be treated by administration of the Toso
CC protein or antibody. Apoptosis related or mediated conditions that can
CC be treated include diseases characterized by T cell overactivity, e.g.
CC Sjogrens connective tissue disorder, autoimmune diseases, diseases where
CC T cells actively destroy cells, including transplant rejection and
CC conditions where cells of any kind that are not dying express Toso
CC appropriately, e.g. cancer of T or B cell origin (where increased
CC apoptosis would be appropriate).

XX Sequence 390 AA;

Query Match 99.6%; Score 2047; DB 20; Length 390;
Best Local Similarity 99.7%; Pred. No. 1.8e-166;
Matches 389; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDRWLWPLYFLPVSGALRLPEVKVEGELGSGVTIKCPLPEMHVRIYLCREMGAGSGTCT 60
|||
DB 1 MDFWLWPLYFLPVSGALRLPEVKVEGELGSGVTIKCPLPEMHVRIYLCREMGAGSGTCT 60

QY 61 VYSTNFIKAEYKGRVTLKQYPRKNLFLVEVTOLTESDSGVYACGAGMNTDRGKTOKVTL 120
|||
DB 61 VYSTNFIKAEYKGRVTLKQYPRKNLFLVEVTOLTESDSGVYACGAGMNTDRGKTOKVTL 120

QY 121 NVHSEYEPSWEQPMPTPKWFHLPYLFQMPAYASSSKFVTTPAQRGKVPVHHSSP 180
|||
DB 121 NVHSEYEPSWEQPMPTPKWFHLPYLFQMPAYASSSKFVTTPAQRGKVPVHHSSP 180

QY 181 TQOITHRPRVSRASSVAGDKPRTFLPSTTASKISALEGLLKQPTPSYNNHTRLHRQALD 240
|||
DB 181 TQOITHRPRVSRASSVAGDKPRTFLPSTTASKISALEGLLKQPTPSYNNHTRLHRQALD 240

QY 241 YGSGSGREGQGFHILPTILGLFLALLGLVVKRAVERRKALSRRLAVRRALESSQ 300
|||
DB 241 YGSGSGREGQGFHILPTILGLFLALLGLVVKRAVERRKALSRRLAVRRALESSQ 300

QY 301 RPRGSPRPRSONNIYSACPRRARGADAAGTGEAPVPGAPLPAPLQVSESPWLHAPSL 360
|||
DB 301 RPRGSPRPRSONNIYSACPRRARGADAAGTGEAPVPGAPLPAPLQVSESPWLHAPSL 360

QY 361 KTSCEYVSLYHQPAAAMMEDSDSDYINVPA 390
ID AAY05001
Db 361 KTSCEYVSLYHQPAAAMMEDSDSDYINVPA 390

RESULT 3
AAAY05001
ID AAY05001 standard; Protein; 390 AA.
AC AAY05001;
XX
XX
XX
DT 16-JUN-1999 (first entry)
XX
DE Human PIGRL-1 protein sequence.
XX
XX
KW PIGRL-1; human; autoimmune disease; hyper-IgM Immunodeficiency; HIM;
KW X-linked Severe Combined Immunodeficiency; XSCID; IgA deficiency;
KW diagnosis; therapy.
XX
XX
OS Homo sapiens.
XX
XX
PN EP905238-A2.
XX
XX
PD 31-MAR-1999.
XX
XX
XX 14-AUG-1998; 98EP-0306487.
XX
XX 30-OCT-1997; 97US-0961564.
PR 25-AUG-1997; 97US-0056935.
XX
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
XX
XX
XX Sweet RW, Truneh A, Wu S;
PI
XX
XX WPI; 1999-192666/17.
DR N-PSDB; AAX281178.
XX
XX

New polypeptides encoding PIGRL-1 useful for treating diseases such as X-linked Severe Combined Immunodeficiency
Claim 11; Page 7; 26pp; English.

This sequence is the human PIGRL-1 protein of the invention.
Autoimmune diseases involving altered expression or activity of PIGRL-1 may include Hyper-IgM Immunodeficiency (HIM), X-linked Severe Combined Immunodeficiency (XSCID) and IgA deficiency. These diseases can be diagnosed or susceptibility to them predicted by: (1) determining whether there is a mutation in the genomic copy of the gene encoding PIGRL-1; or (2) measuring the amount of PIGRL-1 in a sample derived from the patient. Patients deficient in PIGRL-1 can be treated by administering either the PIGRL-1 DNA or its complement or an agonist of PIGRL-1 to the patient. Patients with excessive expression or activity of PIGRL-1 can be treated by administering an antagonist of PIGRL-1, an antisense nucleic acid molecule which inhibits the expression of PIGRL-1 or administering sufficient PIGRL-1 to compete with the endogenous activity. PIGRL-1 can be used to identify its agonists by contacting a cell expressing PIGRL-1 with a candidate compound in the presence of a signal system and noting the candidate as an agonist if a signal is produced. The same method can be used to identify antagonists of PIGRL-1 but the presence of an antagonist is indicated by a decrease in production of the signal. Antibodies against PIGRL-1 may be used to isolate or identify clones expressing PIGRL-1. Polynucleotides encoding PIGRL-1 may be used to identify chromosomal mutations in the gene encoding PIGRL-1 in patients. This information may then be correlated with the incidence of autoimmune disease in those patients to identify whether the mutation causes the disease.

XX Sequence 390 AA;

Query Match 99.6%; Score 2047; DB 20; Length 390;
Best Local Similarity 99.7%; Pred. No. 1.8e-166;
Matches 389; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDRWLPLFLPVSGALRILPEVKVEGELGGSVTIKCPLPEMHVRIYILCREMAGSGTCGT 60
Db 1 MDRWLPLFLPVSGALRILPEVKVEGELGGSVTIKCPLPEMHVRIYILCREMAGSGTCGT 60
QY 61 VVSTTNFKAKEYKGRVTLKQYPRKNLFLVEVTQLTESDSGVYACGAGMNTDRGKTQKVTL 120
Db 61 VVSTTNFKAKEYKGRVTLKQYPRKNLFLVEVTQLTESDSGVYACGAGMNTDRGKTQKVTL 120
QY 121 NVHSEYEPSWEQPMPEPKWFHLPYLQMPAYASSKSFVTRVTPAQRKGVPVHHSSP 180
Db 121 NVHSEYEPSWEQPMPEPKWFHLPYLQMPAYASSKSFVTRVTPAQRKGVPVHHSSP 180
QY 181 TTQITHRPRVSRASSVAGDKPRTPELPTSTASKISALEGLLKQTPSYNHHTLHQRALD 240
Db 181 TTQITHRPRVSRASSVAGDKPRTPELPTSTASKISALEGLLKQTPSYNHHTLHQRALD 240
QY 241 YGSQSGREGQGFHILPTITLGLFLALLGLVYKRAVERRKALSRARLAVMRALSSQ 300
Db 241 YGSQSGREGQGFHILPTITLGLFLALLGLVYKRAVERRKALSRARLAVMRALSSQ 300
QY 301 RPRGSPRPRSQNNIYSACPRRARGADAAGTGEAPVPGCAPLPAPLOVSESPWLHAPSL 360
Db 301 RPRGSPRPRSQNNIYSACPRRARGADAAGTGEAPVPGCAPLPAPLOVSESPWLHAPSL 360
QY 361 KTSCEYVSLYHQPAAAMMEDSDSDYINVPA 390
Db 361 KTSCEYVSLYHQPAAAMMEDSDSDYINVPA 390

RESULT 4
AAE05349

ID AAE05349 standard; Protein; 422 AA.

XX AC AAE05349;

DT 12-SEP-2001 (first entry)

DE Mouse Toso protein.

KW Mouse; cytostatic; antiinflammatory; immunoregulatory; tissue integrity;
KW wound healing; immune response; vaccine; cancer; asthma; allergy;
KW cell trafficking; therapy; secreted protein; Fas-induced apoptosis;
KW Toso.

XX OS Mus sp.

XX PN WO200148192-A1.

XX PD 05-JUL-2001.

XX PF 21-DEC-2000; 2000WO-NZ00256.

XX PR 23-DEC-1999; 99US-0171678.

XX PR 28-NOV-2000; 2000US-0724864.

XX PA (GENE-) GENESIS RES & DEV CORP LTD.

XX PI Watson JD, Murison JG;

XX DR WPI; 2001-425665/45.

XX DR N-PSDB; AAD10117.

XX Novel isolated polypeptide useful to isolate corresponding interacting proteins or other compounds, to quantitatively determine levels of interacting proteins or other compounds, and as therapeutic target -

PS Claim 6; Page 78-79; 101pp; English.

XX The patent discloses novel polynucleotides and their corresponding proteins which play a major role in induction of growth, cell migration and proliferation, cell-cell interaction and the differentiation of tissue-specific cells. These proteins are important in the maintenance of tissue integrity and thus are important in wound healing. They are

QY 10 FLVPVSGALRILPEVKVEGELGSGVTIKCPLPE-----MHVRIYLCREMAGSGTCGTIVVSTT 65
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 Db 13 FSGVSTKSPIFGPQEVSSIEGDSVSYCYYPDTSVNRHTRKYWCROGA-SGMCTTLLSSN 71
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 QY 66 NFIAEYKGRVILKQYPRKNLFLVEVTQLTESDSGVYACAGMNTDRGKTQKVLNVHSE 125
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 Db 72 GYLSKEYSGRANLINFENNTFVINEQLTQDDTGSYKCGLG-TSNRGLSFDVSLEV--- 127
 : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 QY 126 YEPSWEQPMPTPKWPHLPLFQMPAYASSSKFVTRVTPAQRGKVP 173
 : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 Db 128 -----SQVPELPSDTHV-----YTKDIGRNVITIECPFKRENVP 160
 : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 RESULT 8
 AAG65698
 ID AAG65698 standard; protein; 771 AA.
 AC AAG65698;
 DT 07-JAN-2002 (first entry)
 XX Mouse polymeric immunoglobulin receptor (pIgR) sequence.
 DE
 XX
 XX Polymeric immunoglobulin receptor; pIgR; ligand; therapeutic;
 KW carcinoma diagnosis; veterinary; mouse.
 XX
 OS Mus sp.
 XX
 PN W0200172846-A2.
 XX
 PD 04-OCT-2001.
 XX
 XX 26-MAR-2001; 2001WO-US09699.
 XX
 PR 27-MAR-2000; 2000US-192197P.
 PR 27-MAR-2000; 2000US-192198P.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 XX Mostov KE, Chapin SJ, Richman-Eisenstat J;
 PI WPI; 2001-611619/70.
 XX
 XX New ligands binding to a specific region of a polymeric immunoglobulin
 PT receptor, useful for transporting therapeutic or diagnostic
 PT compositions into or across cells expressing pIgR e.g. in drug delivery
 PT
 XX
 PS Disclosure; Fig 1; 102pp; English.
 XX
 CC The invention provides ligands that bind specifically to a region of an
 CC animal cell polymeric immunoglobulin receptor (pIgR). The pIgR cleaves
 CC to produce a stalk region remaining attached to the cell and a secretory
 CC component existing in the organ of interest in several forms. The ligands
 CC do not bind to the stalk or the most abundant form of the secretory
 CC component present in the organ under physiological conditions. The
 CC ligands are useful for transporting therapeutic or diagnostic
 CC compositions into or across cells expressing pIgR, useful to introduce
 CC or transport ligands such as antibodies and/or to deliver biologically
 CC active components such as proteins, nucleic acids or detectable labels.
 CC They are used to deliver therapeutic compositions to mucosal surfaces
 CC such as the gastro-intestinal tract, respiratory system etc. in humans.
 CC They are also useful to label cells expressing pIgR, e.g. to distinguish
 CC epithelial cells from a mixed cell population in pathology studies or to
 CC aid in carcinoma diagnosis (since pIgR expression is reduced in
 CC carcinomas relative to normal epithelium). They can also be used to
 CC deliver veterinary compositions, especially in mammals such as farm,
 CC domestic or wild mammals or birds e.g. birds reared for human
 CC consumption. The present sequence represents a mouse pIgR sequence.
 XX
 XX Sequence 771 AA;
 SQ

Query Match 9.5%; Score 195; DB 22; Length 771;
 Best Local Similarity 34.5%; Pred. No. 6,1e-08;
 Matches 48; Conservative 21; Mismatches 54; Indels 16; Gaps 4;
 QY 10 FLVPVSGALRILPEVKVEGELGSGVTIKCPLPE-----MHVRIYLCREMAGSGTCGTIVVSTT 65
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 Db 13 FSGVSTKSPIFGPQEVSSIEGDSVSYCYYPDTSVNRHTRKYWCROGA-SGMCTTLLSSN 71
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 QY 66 NFIAEYKGRVILKQYPRKNLFLVEVTQLTESDSGVYACAGMNTDRGKTQKVLNVHSE 125
 | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 Db 72 GYLSKEYSGRANLINFENNTFVINEQLTQDDTGSYKCGLG-TSNRGLSFDVSLEV--- 127
 : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 QY 126 YEPSWEQPMPTPKWPHL 144
 : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 Db 128 -----SQVPELPSDTHV 139
 : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
 RESULT 9
 AAW03181
 ID AAW03181 standard; Protein; 769 AA.
 AC AAW03181;
 DT 24-FEB-1997 (first entry)
 XX Rat poly-immunoglobulin receptor.
 DE
 XX
 KW Rat; immunoglobulin; receptor; protection protein; mutants;
 KW heavy chain; antigen binding domain; protection; pathogen;
 KW mucosal; environment; gastrointestinal; passive; immunisation;
 KW Guy's 13 antibody; prevention; dental caries; Streptococcus;
 KW poly; sorbinus.
 XX
 OS Rattus rattus.
 XX
 FH Key Location/Qualifiers
 FT Region 13..45
 FT /note= "putative immunoglobulin binding residues
 FT Of domain I"
 FT Domain 1..120
 FT /label= domain_I
 FT Domain 110..230
 FT /label= domain_II
 FT Domain 210..340
 FT /label= domain_III
 FT Domain 320..450
 FT /label= domain_IV
 FT Domain 440..550
 FT /label= domain_V
 FT Region 550..606
 FT /note= "external portions of domain VI"
 FT Region 550..627
 FT /note= "external portions of domain VI"
 FT Region 625..660
 FT /label= transmembrane_segment
 FT Region 650..769
 FT /label= intracellular_portion
 FT
 XX W09621012-A1.
 PN
 XX 11-JUL-1996.
 XX
 XX 27-DEC-1995; 95WO-US16889.
 PR 04-MAY-1995; 95US-0434000.
 PR 30-DEC-1994; 94US-0367395.
 XX
 XX (PLAN-) PLANT BIOTECHNOLOGY INC.
 PA (UNMF-) UNITED MEDICAL & DENTAL SCHOOLS GUYS.
 PA (PLAN-) PLANET BIOTECHNOLOGY INC.
 XX
 PI Hiatt AC, Lehner T, Ma JKC;
 XX


```
DR WPI; 1996-333987/33.
DR N-PSDB; AAT31291.
XX Immunoglobulin and protection protein complex and its prodn. in
XX plants - useful for passive immunisation against mucosal antigens,
PT esp. against S. mutans and S. sorbinus to prevent dental caries
XX
XX Disclosure; Pages 123-127; 152pp; English.
XX
XX The present sequence is the rat poly-immunoglobulin (Ig)
XX receptor, a portion of which corresp. to residues 1-627, pref.
XX 1-606, or esp. residues 13-45, 1-120, 110-230, 210-340, 320-450,
XX 440-550, 550-606 or 550-627 comprises a protection protein (PP).
XX The Ig of the invention comprises a PP as above in association with
XX an Ig derived heavy chain, having at least a portion of an antigen
XX (Ag) binding domain. The PP protects the Ig in harsh mucosal, e.g.
XX gastrointestinal, environments, therefore enhancing its
XX effectiveness in passively immunising animals against mucosal
XX pathogens. The Ag binding domain is specifically derived from the
XX Guy's 13 antibody, and the Ig can be used to prevent dental caries
XX by binding, e.g. Streptococcus mutans serotypes c, e and f, or
XX S. sorbinus serotypes d and g.
XX
XX Sequence 769 AA;
SQ
Query Match 9.0%; Score 185; DB 17; Length 769;
Best Local Similarity 34.5%; Pred. No. 4.4e-07;
Matches 41; Conservative 21; Mismatches 41; Indels 16; Gaps 4;
QY 30 GGSVTIKCPLE----MHVRIYLCREMAGSGTCGTVVSTNFIKAEYKGRVTLKQYPRKN 85
DB 33 GNSVSTICYPTSVNRHTRKYWCROGA-NGYCATLISSNGYLSKEYSGRASLINFPPNS 91
QY 86 LFLVEVTQLTESDGVYACGAGMNTDRGKTQKVTNLNVHSEYEPSEWEEQMPETPKWFHL 144
DB 92 TFINIAHLTQEDTGSYKCLG-TTNRGLFDVSLV-----SQVPEPNDTHV 139
RESULT 10
AAG5697
ID AAG65697 standard; protein; 769 AA.
AC AAG65697;
XX
XX 07-JAN-2002 (first entry)
XX
XX Rat polymeric immunoglobulin receptor (pIgR) sequence.
XX
XX Polymeric immunoglobulin receptor; pIgR; ligand; therapeutic;
XX carcinoma diagnosis; veterinary; rat.
XX
XX Rattus sp.
XX
XX WO200172846-A2.
XX
XX 04-OCT-2001.
XX
XX 26-MAR-2001; 2001WO-US09699.
XX
XX 27-MAR-2000; 2000US-192197P.
XX
XX 27-MAR-2000; 2000US-192198P.
XX
XX (REGC ) UNIV CALIFORNIA.
XX
XX Mostov KE, Chapin SJ, Richman-Eisenstat J;
XX
XX WPI; 2001-611619/70.
XX
XX New ligands binding to a specific region of a polymeric immunoglobulin
XX receptor, useful for transporting therapeutic or diagnostic
XX compositions into or across cells expressing pIgR e.g. in drug delivery
XX
```

```
PS Disclosure; Fig 1; 102pp; English.
XX
XX The invention provides ligands that bind specifically to a region of an
XX animal cell polymeric immunoglobulin receptor (pIgR). The pIgR cleaves
XX to produce a stalk region remaining attached to the cell and a secretory
XX component existing in the organ of interest in several forms. The ligands
XX do not bind to the stalk or the most abundant form of the secretory
XX component present in the organ under physiological conditions. The
XX ligands are useful for transporting therapeutic or diagnostic
XX compositions into or across cells expressing pIgR, useful to introduce
XX or transport ligands such as antibodies, nucleic acids or detectable labels.
XX active components such as proteins, nucleic acids or detectable labels.
XX They are used to deliver therapeutic compositions to mucosal surfaces.
XX such as the gastro-intestinal tract, respiratory system etc. in humans.
XX They are also useful to label cells expressing pIgR, e.g. to distinguish
XX epithelial cells from a mixed cell population in pathology studies or to
XX aid in carcinoma diagnosis (since pIgR expression is reduced in
XX carcinomas relative to normal epithelium). They can also be used to
XX deliver veterinary compositions, especially in mammals such as farm,
XX domestic or wild mammals or birds e.g. birds reared for human
XX consumption. The present sequence represents a rat pIgR sequence.
XX
XX Sequence 769 AA;
SQ
Query Match 9.0%; Score 185; DB 22; Length 769;
Best Local Similarity 34.5%; Pred. No. 4.4e-07;
Matches 41; Conservative 21; Mismatches 41; Indels 16; Gaps 4;
QY 30 GGSVTIKCPLE----MHVRIYLCREMAGSGTCGTVVSTNFIKAEYKGRVTLKQYPRKN 85
DB 33 GNSVSTICYPTSVNRHTRKYWCROGA-NGYCATLISSNGYLSKEYSGRASLINFPPNS 91
QY 86 LFLVEVTQLTESDGVYACGAGMNTDRGKTQKVTNLNVHSEYEPSEWEEQMPETPKWFHL 144
DB 92 TFINIAHLTQEDTGSYKCLG-TTNRGLFDVSLV-----SQVPEPNDTHV 139
RESULT 11
AAW50033
ID AAW50033 standard; Protein; 532 AA.
XX
XX AAW50033;
XX
XX 26-JUN-1998 (first entry)
XX
XX Human immunity related factor.
XX
XX Lymph node; human; immunity related factor; research; treatment;
XX immune disease; infectious disease.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 1..16
XX /label= sig_peptide
XX Peptide 17..532
XX /label= mat_peptide
XX
XX JP10072495-A.
XX
XX 17-MAR-1998.
XX
XX 11-JUN-1997; 97JP-0153218.
XX
XX 13-JUN-1996; 96JP-0152362.
XX
XX (ASAH ) ASahi KASEI KOGYO KK.
XX
XX WPI; 1998-234766/21.
XX N-PSDB; AAV20383.
XX
XX Immunity related factor - useful in the treatment of immune related
XX and infectious diseases
PT
```


CC immunoglobulin A (sIgA). It can be used as part of a method for
CC the production of sig molecules. This method is useful for
CC producing commercial quantities of sig (especially sIgA) to treat
CC or prevent infections. In particular, sIgA produced by the method
CC can be used to prevent or treat infections in mammals, birds or
CC fish; especially systemic infections or infections at a mucosal
CC surface. It is especially useful to prevent or treat infection
CC with human immunodeficiency virus (HIV), respiratory syncytial
CC virus, flu virus or cold virus. The method allows production of
CC commercial quantities of sig molecules for therapeutic use, not
CC previously possible; production using non-plant cells and a
CC single cell type is more efficient than a previous multi-step
CC process of fusing recombinant plant cells, and avoids alterations
CC of the sig by plant cells. sIgA molecules are more stable
CC and resistant to proteolysis than previously used IgA molecules,
CC and can be administered to prevent as well as to treat infections,
CC unlike e.g. IgG and IgM molecules.

XX Sequence 607 AA;

Query Match 8.9%; Score 182; DB 20; Length 607;
Best Local Similarity 37.4%; Pred. No. 5.7e-07;
Matches 43; Conservative 20; Mismatches 40; Indels 12; Gaps 5;

QY 12 PVSGALRLPEVKVEGELGSGVTIKCLPLP-----EMHVRVYLCREMAGSGTGTWVSTTNF 67
DB 21 PIFG-----PE-EVNSVEGNSVITCYPTSVNRHTRKYWCROGARGG-CITLISSEGY 73
QY 68 IKAERYGRVTLKOYPRKNLFLVEVTLTSDSGVYACGAGMNTDRGKTOKVTLVN 122
DB 74 VSSKYAGRANLTNFPENGTFVNVIAQLSQDSDSGRYKCGLGINS-RGLSFDVSLEV 127

RESULT 14

AAW03178
ID AAW03178 standard; Protein; 746 AA.

XX AC AAW03178;

DT 24-FEB-1997 (first entry)

DE Human poly-immunoglobulin receptor.

XX Human; immunoglobulin; receptor; protection protein; mutants;
KW heavy chain; antigen binding domain; protection; pathogen;
KW mucosal; environment; gastrointestinal; passive; immunisation;
KW Guy's 13 antibody; prevention; dental caries; Streptococcus;
KW poly; sorbinus.

XX OS Homo sapiens.

XX Key Location/Qualifiers
FH Region 13..45
FT /note= "putative immunoglobulin binding residues
FT of domain I"

FT Domain 1..120

FT /label= domain_I

FT Domain 110..230

FT /label= domain_II

FT Domain 210..340

FT /label= domain_III

FT Domain 320..450

FT /label= domain_IV

FT Domain 440..550

FT /label= domain_V

FT Region 550..606

FT /note= "external portions of domain VI"

FT Region 550..627

FT /note= "external portions of domain VI"

FT Region 625..660

FT /label= transmembrane_segment

FT Region 650..746

FT /label= intracellular_portion

XX WO9621012-A1.
XX 11-JUL-1996.
XX 27-DEC-1995; 95WO-US16889.
XX 04-MAY-1995; 95US-0434000.
XX 30-DEC-1994; 94US-0367395.
XX (PLAN-) PLANT BIOTECHNOLOGY INC.
XX (UNME-) UNITED MEDICAL & DENTAL SCHOOLS GUYS.
XX (PLAN-) PLANET BIOTECHNOLOGY INC.
XX Hiatt AC, Lehner T, Ma JKC;
XX WPI; 1996-333987/33.
XX N-PSDB; AAT31288.
XX Immunoglobulin and protection protein complex and its prodn. in
XX plants - useful for passive immunisation against mucosal antigens,
XX esp. against S. mutans and S. sorbinus to prevent dental caries
XX Disclosure; Pages 105-108; 152pp; English.
XX The present sequence is the human poly-immunoglobulin (Ig)
XX receptor, a portion of which corresp. to residues 1-627, pref.
XX 1-606, or esp. residues 13-45, 1-120, 110-230, 210-340, 320-450,
XX 440-550, 550-606 or 550-627 comprises a protection protein (PP).
XX The Ig of the invention comprises a PP as above in association with
XX an Ig derived heavy chain, having at least a portion of an antigen
XX (Ag) binding domain. The PP protects the Ig in harsh mucosal, e.g.
XX gastrointestinal, environments, therefore enhancing its
XX effectiveness in passively immunising animals against mucosal
XX pathogens. The Ag binding domain is specifically derived from the
XX Guy's 13 antibody, and the Ig can be used to prevent dental caries
XX by binding, e.g. Streptococcus mutans serotypes c, e and f, or
XX S. sorbinus serotypes d and g.

XX Sequence 746 AA;

Query Match 8.9%; Score 182; DB 17; Length 746;
Best Local Similarity 37.4%; Pred. No. 7.5e-07;
Matches 43; Conservative 20; Mismatches 40; Indels 12; Gaps 5;

QY 12 PVSGALRLPEVKVEGELGSGVTIKCLPLP-----EMHVRVYLCREMAGSGTGTWVSTTNF 67
DB 3 PIFG-----PE-EVNSVEGNSVITCYPTSVNRHTRKYWCROGARGG-CITLISSEGY 55

QY 68 IKAERYGRVTLKOYPRKNLFLVEVTLTSDSGVYACGAGMNTDRGKTOKVTLVN 122
DB 56 VSSKYAGRANLTNFPENGTFVNVIAQLSQDSDSGRYKCGLGINS-RGLSFDVSLEV 109

RESULT 15

AAW03179

ID AAW03179 standard; Protein; 757 AA.

XX AC AAW03179;

XX 24-FEB-1997 (first entry)

XX Bovine poly-immunoglobulin receptor.

DE Bovine; immunoglobulin; receptor; protection protein; mutants;
KW heavy chain; antigen binding domain; protection; pathogen;
KW mucosal; environment; gastrointestinal; passive; immunisation;
KW Guy's 13 antibody; prevention; dental caries; Streptococcus;
KW poly; sorbinus; cow.

XX OS Bos taurus.

XX Key Location/Qualifiers

```

FT Region 13..45
FT /note= "putative immunoglobulin binding residues
FT of domain I"
FT
FT Domain 1..120
FT /label= domain_I
FT Domain 110..230
FT /label= domain_II
FT Domain 210..340
FT /label= domain_III
FT Domain 320..450
FT /label= domain_IV
FT Domain 440..550
FT /label= domain_V
FT Region 550..606
FT /note= "external portions of domain VI"
FT Region 550..627
FT /note= "external portions of domain VI"
FT Region 625..660
FT /label= transmembrane_segment
FT Region 650..757
FT /label= intracellular_portion
XX
PN W09621012-A1.
XX
XX 11-JUL-1996.
XX
XX 27-DEC-1995; 95WO-US16889.
XX
XX 04-MAY-1995; 95US-0434000.
XX 30-DEC-1994; 94US-0367395.
XX
XX (PLAN-) PLANT BIOTECHNOLOGY INC.
XX (UNNE-) UNITED MEDICAL & DENTAL SCHOOLS GUYS.
XX (PLAN-) PLANET BIOTECHNOLOGY INC.
XX
XX Hiatt AC, Lehner T, Ma JKC;
XX
XX WPI: 1996-333987/33.
XX DR N-PSDB; AAT31289.
XX
XX Immunoglobulin and protection protein complex and its prodn. in
PT plants - useful for passive immunisation against mucosal antigens,
PT esp. against S. mutans and S. sorbinus to prevent dental caries
XX
XX Disclosure; Pages 111-115; 152pp; English.
XX
XX The present sequence is the bovine poly-immunoglobulin (Ig)
XX receptor, a portion of which corresp. to residues 1-627, Pref.
XX 1-606, or esp. residues 13-45, 1-120, 110-230, 210-340, 320-450,
XX 440-550, 550-606 or 550-627 comprises a protection protein (PP).
XX The Ig of the invention comprises a PP as above in association with
XX an Ig derived heavy chain, having at least a portion of an antigen
XX (Ag) binding domain. The PP protects the Ig in harsh mucosal, e.g.
XX gastrointestinal, environments, therefore enhancing its
XX effectiveness in passively immunising animals against mucosal
XX pathogens. The Ag binding domain is specifically derived from the
XX Guy's 13 antibody, and the Ig can be used to prevent dental caries
XX by binding, e.g. Streptococcus mutans serotypes c, e and f, or
XX S. sorbinus serotypes d and g.
XX
XX Sequence 757 AA;

Query Match 8.9%; Score 182; DB 17; Length 757;
Best Local Similarity 34.1%; Pred. No. 7.7e-07;
Matches 42; Conservative 21; Mismatches 44; Indels 16; Gaps 4;

QY 17 LRILPVKVEGEL-----GGSVTKCPLP-----EMHVRVYLCREMAGSGTCTGVV 62
DB 10 LAIFPVVSMKSFICPEEVSVEGRSVKCYPPPTSVNHTKRYWCRQGA-QGRCTTLL 68
QY 63 STNFTKAEYKGRVTLKOYPRKNLFLVEVTLTSDSGVYACGAGMNTDRGKTOKVTLNV 122
DB 69 SSGEYVSDDYVGRANLTNFPESCTFVDISHLTHKDSGRYKCGLGISS-RGLNFDVSLEY 127

```

QY 123 HSE 125
 Db 128 SQD 130

Search completed: October 28, 2002, 17:30:51
 Job time : 44.8871 secs